DOCUMENT RESUME

ED 076 225 LI 004 338

AUTHOR	Fox, Ann M.
TITLE	The Amenability of a Cataloging Process to Simulation by Automatic Techniques. Final Report.
INSTITUTION	Illinois Univ., Urbana. Graduate School of Library Science.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-8-E-123
PUB DATE	May 72
GRANT	OEG-5-9-235123-0019(095)
NOTE	477p.; (42 References); Doctoral Thesis; University of Illinois, Urbana-Champaign
AVAILABLE FROM	University Microfilms, 200 N. Zeeb Road, Ann Arbor, Michigan 48106 (UM-73-9934, HC $10.00, MF $4.00)
EDRS PRICE	MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS	*Automation; *Cataloging; Doctoral Theses; Graphs; Library Technical Processes; Models; *Simulation

ABSTRACT

This study attempts to determine whether the human intellectual process of cataloging bibliographic materials, using the Anglo-American (AA-1967) and American Library Association (ALA-1949) cataloging codes, can be simulated by automatic techniques. The specific cataloging process is that which concerns selection of entry. Automatic techniques used to simulate the entry process are directed tree graphs. It is the basic assumption of the study that the AA rules of entry are more amenable to simulation than the ALA rules of entry. Measurement of amenability was determined in terms of four criterion variables. With the exception of one variable, measurement did not support this assumption. The directed tree graphs for the two sets of entry rules were tested for their accuracy to assign correct main entries, and their ability to accomplish this assignment automatically. Both tests showed the graphs for the two sets of rules to be highly and equally, or almost equally, accurate and automatic. Findings of the study apply to human application of the graphs. The economic or technical feasibility of the graphs is not considered. (Author)

ED 076225

FILMED FROM BEST AVAILABLE COPY

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY
Ann M. Fox
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

Final Report

Project No. 8 E-123
Grant No. OEG 5-9 235123 0019(095)

Ann M. Fox
University of Illinois
Urbana, Illinois 61820

THE AMENABILITY OF A CATALOGING PROCESS TO SIMULATION BY AUTOMATIC TECHNIQUES

May 1972

U. S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of Education

Bureau of Research

LI 004 338

© Copyright by
Ann Sandberg Fox
1972

THE AMENABILITY OF A CATALOGING PROCESS TO SIMULATION BY AUTOMATIC TECHNIQUES

BY

ANN SANDBERG FOX

A.B., College of Our Lady of the Elms, 1953
M.S., Catholic University of America, 1955

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Library Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1972

Urbana, Illinois

THE AMENABILITY OF A CATALOGING PROCESS TO SIMULATION BY AUTOMATIC TECHNIQUES

Ann Sandberg Fox, Ph.D.
Graduate School of Library Science
University of Illinois at Urbana-Champaign, 1972

This study attempts to determine whether the human intellectual process of cataloging bibliographic materials, using the AA (1967) and ALA (1949) cataloging codes, can be simulated by automatic techniques. The specific cataloging process is that which concerns selection of entry. Automatic techniques used to simulate the entry process are directed tree graphs.

It is the basic assumption of the study that the AA rules of entry are more amenable to simulation than the ALA rules of entry. Measurement of amenability was determined in terms of four criterion variables. With the exception of one variable, measurement did not support this assumption.

The directed tree graphs for the two sets of entry rules were tested for their accuracy to assign correct main entries, and their ability to accomplish this assignment automatically. Both tests showed the graphs for the two sets of rules to be highly and equally, or almost equally, accurate and automatic.

Findings of the study apply to human application of the graphs. The economic or technical feasibility of the graphs is not considered.

ACKNOWLEDGMENT

The undertaking and completion of a study of this nature reflects the interest and cooperation of a great many people. During the entire period of this research, from the formulation of the research proposal to the reading of the final manuscript, the investigator has had the benefit of very frequent and lengthy discussions with her advisor, Dr. Herbert Goldhor. His advice and, in particular, his encouragement through difficult periods of the research are especially appreciated. The lengthy discussions with Mrs. Kathryn L. Henderson, who read this manuscript in progressive stages and offered valuable criticisms and suggestions on points of content, are also appreciated.

At various stages of this study, the investigator has had the benefit of many comments and suggestions from members of the cataloging staff at the University of Illinois library and the faculties of the Graduate School of Library Science and the Department of Linguistics. Special thanks are given to the members of the doctoral committee, in addition to Dr. Herbert Goldhor and Mrs. Kathryn L. Henderson, Dr. Henry R. Kahane, Dr. Rolland E. Stevens, and Miss Betty Croft.

Thanks are also given to the participants in the experiment, Mrs. Inez Wildemuth, for her expert clerical assistance, and Miss Ruth I. Jacobson, for her execution of the graphs.

In very obvious ways, the course of this research has been influenced strongly by the work of Noam Chomsky, the teaching of

Dr. Robert B. Lees and Dr. Henry R. Kahane, the philosophical insights of Mr. Leonard Jolley, and the criticisms and suggestions provided by her husband and linguist, Dr. Robert P. Fox.

Acknowledgment is also given to the American Library Association for permitting certain of the cataloging rules in the ALA (1949) and Anglo-American (1967) codes to be reprinted in the appendices of the study.

Work on the formulation of the proposal was begun when the investigator was a University of Illinois Fellow in the Graduate School of Library Science, and was carried out with support extended to the University of Illinois by the United States Office of Education under Contract No. OEG-5-9-235123-0019(095), Project No. 8-E-123.

TABLE OF CONTENTS

CHAPTER Page

I THE PROBLEM AND RELATED LITERATURE 1

Related Literature 4
Summary 10

II HYPOTHESIS AND GENERAL RESEARCH METHOD 12

The Hypothesis 12
General Research Method 19
Summary 24

III RULES OF ENTRY SELECTION AND AMENABILITY 26

Selection of Rules 26
Measurement of Amenability 49
Findings 75
Summary 101

IV FORMULATION OF THE SIMULATION MODELS 103

Formulation of the Characterization and Design of the Simulation Models 103
The Simulation Models 162

V FIRST TEST OF ADEQUACY 166

Procedures 166
Findings 183

VI SECOND TEST OF ADEQUACY 246

Procedures 246
Findings 260

VII SUMMARY AND CONCLUSIONS 292

Summary of the Findings 292
Limitations and Assumptions 299
Conclusions 300
Practical Implications 303
Suggestions for Further Research 305

Page

BIBLIOGRAPHY . . . 308

APPENDIX

A RULES OF ENTRY SELECTED FOR CHARACTERIZATION . . . 311
B EXPERIMENTAL RULES . . . 328
C KEY TO SYMBOLS AND TERMS USED IN THE EXPERIMENTAL RULES . . . 372
D INSTRUCTIONS TO PARTICIPANTS AND SAMPLE TITLE PAGE . . . 375
E TITLE PAGES USED IN THE EXPERIMENT . . . 415
F QUESTIONNAIRE TO PARTICIPANTS AND CORRESPONDENCE . . . 460

VITA . . . 464

LIST OF TABLES

Table Page

1 Frequency of Use of AA Rules of Entry 31

2 Frequency of Use of ALA Rules of Entry 33

3 Frequency of Main Entry Type 37

4 Frequency of Main Entry Type for Shipments Comprising AA and ALA Samples 38

5 Correlation of AA Rules of Entry with ALA Rules of Entry from Tables 1 and 2 40

6 Corpus of AA Rules of Entry and Comparable ALA Rules of Entry Selected for Characterization 50

7 Structural Descriptions for the Corpus of AA Rules of Entry 59

8 Structural Descriptions for the Corpus of ALA Rules of Entry 65

9 Redundancy in the Corpus of AA Rules of Entry as Shown in the Structural Descriptions 77

10 Redundancy in the Corpus of ALA Rules of Entry as Shown in the Structural Descriptions 78

11 Complexity in the Corpus of AA Rules of Entry as Shown in the Number of Conditions and Their Frequency in the Structural Descriptions 81

12 Complexity in the Corpus of ALA Rules of Entry as Shown in the Number of Conditions and Their Frequency in the Structural Descriptions 85

13 Specificity in the Conditions of the Structural Descriptions for the Corpus of AA Rules of Entry . . . 91

14 Specificity in the Conditions of the Structural Descriptions for the Corpus of ALA Rules of Entry . . . 95

15 Distribution of Samples by AA and Comparable ALA Rules of Entry 171

Table Page

16 Distribution of AA and ALA Samples by Main Entry Type 174

17 Entries Paired by Sets of AA and Comparable ALA Rules of Entry for Main Entry Type 176

18 Recognitions Produced by Application of AA and ALA Recognition Rules to Library of Congress English Language Proof Slips 185

19 Distribution of AA and ALA Samples by Recognition Rules 186

20 Type of Unsuccessful AA and ALA Recognitions 188

21 Zero Recognitions Produced by Recognition Rules 190

22 Erroneous Recognitions Produced by Recognition Rules . . 195

23 Main Entry Selections Produced by Application of AA and ALA Experimental Selection Rules to Recognitions . . 199

24 Comparison of Author and Title Main Entries Produced by AA and ALA Experimental Selection Rules with Corresponding Rules of Entry 199

25 Performance of AA Experimental Selection Rules as Applied to AA Recognitions 203

26 Performance of ALA Experimental Selection Rules as Applied to ALA Recognitions 204

27 Performance of Similar AA and ALA Experimental Selection Rules 206

28 Correspondence of Main Entries Produced by AA Experimental Selection Rules with AA Rules of Entry 208

29 Correspondence of Main Entries Produced by ALA Experimental Selection Rules of ALA Rules of Entry 209

30 Main Entries Correlated by AA Rules of Entry and Comparable ALA Rules of Entry 212

31 Number of Incorrect AA and ALA Main Entry Selections Produced by Unsuccessful Recognitions 217

Table Page

32 Number and Type of Main Entry Errors Produced by Incomplete Characterization 221

33 Number and Type of Main Entry Errors Produced by Deficient Characterization 225

34 Complexity in the Models as Measured by Number of Rules, Nodes, and Steps for Selecting Main Entry . . . 236

35 Redundancy in the Verbal Rules as Observed by the Number of Reduced Rules in the Models 236

36 Complexity in the Verbal Rules as Observed by the Number of Nodes in the Experimental Rules 242

37 Specificity in the Verbal Rules as Observed by the Number of Conditions Amenable to Formal Characterization 242

38 Sample of Title Pages Correlated by Corpus of AA and Comparable ALA Rules of Entry 251

39 Number of Title Pages in Each Subset with Correct Application of Recognition Rules 262

40 Title Pages and AA Recognition Rules Incorrectly Applied 264

41 Title Pages and ALA Recognition Rules Incorrectly Applied 264

42 Types of Errors Accounting for Misapplication of AA Recognition Rules 266

43 Types of Errors Accounting for Misapplication of ALA Recognition Rules 266

44 Number of Title Pages in Each Subset with Correct Application of Experimental Selection Rules 273

45 Application of Experimental Selection Rules to Paired Subsets of Title Pages 276

46 Source of Errors Accounting for Misapplication of AA Experimental Selection Rules 279

Table Page

47 Source of Errors Accounting for Misapplication of ALA Experimental Selection Rules 279

48 Hours Taken Per Model 286

49 Responses to Questionnaire 288

CHAPTER I

THE PROBLEM AND RELATED LITERATURE

One of the most critical problems confronting the library profession is that of accelerating the processing of library materials. A persistent problem, the situation has been particularly aggravated by the prodigious rate of increase in the number and kinds of publications acquired by libraries during the last decades.

A cause of serious delay in the processing and, in turn, the availability of bibliographic materials is the demanding and time-consuming task of cataloging them. In response to the explosive growth of acquisitions and the increasing needs and demands of readers, libraries have sought to accelerate the supply of cataloging information through mechanizing certain of their cataloging procedures, and through increased reliance on centralized cataloging services. Developments in these areas have resulted in a rapid improvement of many of the libraries' operational techniques that are concerned with the production and distribution of cataloging copy.

As vital as are these improvements, the central problem remains. The supply of cataloging information is initially and critically dependent upon the availability of personnel to provide it. The enormous increase in acquisitions has required the assignment of more personnel than are presently or in the future likely to be available.

In addition to the shortage of qualified library manpower, another impediment to expediting cataloging copy quickly is the time-consuming character of the cataloging codes. Both the number and variety of rules that require human consultation defy expediency in cataloging. In the report of the cataloging-in-source experiment, catalogers at the Library of Congress spent an average of 38 minutes per entry.[1] More recently, cataloging time for high-priority titles received at the Library of Congress has been eleven days. For other books, it has averaged thirty-five and one-half days.[2] In a very practical sense, automation of this process would fulfill a vital function in saving a considerable amount of qualified human cataloger time and thereby facilitate more, quicker, and perhaps (after some time) cheaper cataloging copy.

In this connection, the question arises as to whether the human cataloging process can be simulated by machine or automatic techniques. Simulating the cataloging process is essentially a mechanization of the intellectual operations of the human cataloger. The generally conceded complexity of the process and the abstract and verbal inarticulateness of the cataloging codes are considered to be serious obstacles to the simulation of this task. It is possible, however, that this process is more complicated linguistically than structurally, and that its basic logical operations are of a kind that could be carried out in computer systems.

---

[1] U. S. Library of Congress, Processing Department, The Cataloging-In-Source Experiment (Washington, D.C.: Library of Congress, 1960), p. 143.

[2] Daniel Melcher, "Cataloging, Processing, and Automation," American Libraries, II (July-August, 1971), 703.

The general problem set for this study was to determine whether the human intellectual process of cataloging bibliographic materials could be simulated by automatic, namely objective, non-intuitive, computer techniques. The specific cataloging process selected for the study was that which concerns selection of entry, or the determination of authorship responsibility. In particular, it formed an attempt to construct a formal theoretical model that would simulate the human process of selecting entry, and to investigate and test its practical application to bibliographic materials.

A precisely constructed formal model of the entry process can serve several purposes. In accomplishing its objective, it can bring to light certain facts about the structure of the entry process, and expose the inadequacies and gaps in the theory underlying the process and the rules available to accomplish it. Similarly, it can reveal limitations in the formulations and design of the model, and expose the source or sources of these limitations. The usefulness of a precise and technical development of the entry process is evident in light of the cataloging situation already mentioned: namely, the overwhelming volume of library acquisitions and the urgency of having them processed rapidly, and the scarcity of cataloging personnel, both of which have indirectly caused log jams in scientific research and public services due to the fact that library materials and information about them are not available at a time when they are most wanted and needed.

Related Literature

The entire history of research on the use of simulation models and techniques for human decision-making and problem-solving tasks spans the period of a decade or, at most, two decades. Compressed within this brief period is an expansive and still burgeoning literature. Since simulation is a young, multidisciplinary area, the literature on its use is unusually scattered and difficult to assemble. Simulation experiments have been reported in such seemingly unrelated areas as language translation, weather analysis and prediction, disease diagnosis, operational or game strategy, business administration, pedagogical instruction, and the taxonomy of archeological artifacts, among others. Studies on simulation methods and model designs comprise still another growing facet of the literature.

Bibliographies and surveys of the literature are available in those fields where the technique has had a great deal of application, notably in education, medicine, engineering, linguistics, military defense, and business administration. Most important of the journals carrying research on its application is Simulation, put out by Simulation Council, Inc., which devotes its pages to research in or directly related to the simulation field. Other journals important in publishing simulation studies are Operational Research Quarterly, Behavioral Science, and The Computer Journal, to mention a few. The major share of research continues to appear in technical reports, the professional journals, and symposium and conference proceedings.

Although the term has eluded precise definition, there is general agreement among research scientists and investigators that simulation is essentially a technique that involves setting up a model of a real situation, and then performing experiments on the model.

Application of this technique to library process has been minimal, and has yet to realize its full potential. Its most recent application has been in the area of library administrative practices and information services.[3] In the few studies reported, simulation techniques were successfully used to analyze library systems and operations in terms of economic, information, and service components.

It is in the areas of indexing, in particular, that the potentialities of simulation have been demonstrated by library and information scientists as well as engineers and linguists. The majority of these studies have been reported in issues of Current Research and Development in Scientific Documentation,[4] and in Automatic Indexing: A State-of-the-Art Report,[5] and LIST 71.[6] An extensive discussion of indexing systems in operation is found in Gerard Salton's Automatic Information Organization and Retrieval,[7] and an overview of developments in the area can be found in the library automation chapters of the Annual Review of Information Science and Technology.[8] A few of the studies, in terms of the similarities in either the objectives or the modeling and experimental design used by the investigators, fall within the framework of the present study and, for that reason, are singled out for brief comment.

---

[3] In particular, see: Norman R. Baker and Richard E. Nance, "The Use of Simulation in Studying Information Storage and Retrieval Systems," American Documentation, XIX (October, 1968), 363-70; Guy J. Caras, "Computer Simulation of a Small Information System," American Documentation, XIX (April, 1968), 120-22; R. L. Chapman, "The Case for Information System Simulation," in Information System Sciences (Washington, D.C.: Spartan Books, Inc., 1964), pp. 477-84.

[4] Current Research and Development in Scientific Information (Washington, D.C.: National Science Foundation, Office of Scientific Information, 1957-69), issues 1-17.

In a series of studies, Phyllis Baxendale, who was one of the first investigators in this area, used computer simulation techniques to determine the capability of the computer to automate the human indexing process.[9] By considering the patterns of scanning that human typically use in indexing she developed procedures for automatic selection of topic sentences, words and phrases. Results of her experiments indicated the advantage of using simulation in the machine selection of an index, and the high percentages of reduction in volume that are possible by all of her techniques without undue loss of content of an article.

---

[5] Mary Elizabeth Stevens, Automatic Indexing: A State-of-the-Art Report (Washington, D.C.: Government Printing Office, 1965), 182 p.

[6] Paul Wasserman, Evelyn Daviel, and Kay Logan, eds., LIST 1971: Library and Information Science Today: An International Registry of Research and Innovation (New York: International, Inc., 1971), 397 p.

[7] Gerard Salton, Automatic Information Organization and Retrieval (New York: McGraw-Hill Book Co., 1968), 514 p.

[8] C. A. Cuadros, ed., Annual Review of Information Science and Technology (Chicago: Encyclopedia Britannica, 1966-70), vols. 1-5.

[9] Phyllis Baxendale, "Machine-Made Index for Technical Literature-An Experiment," IBM Journal of Research and Development, II (October, 1958), 354-61.

Using a different approach, M. J. Ruhl contrasted human concept indexing with machine indexing by comparing the indexing of identical documents appearing in the Chemical Abstracts Subject Index and in the machine prepared Chemical Titles.[10] From this comparison, he showed that more than half of the titles in Chemical Titles included all concepts, or their equivalents, as indexed in the Chemical Abstracts Subject Index.

In an early study on machine indexing and text searching by Don R. Swanson, human indexing and retrieval were pitted against machine retrieval.[11] An assumption of this study was that if the computer had full texts available for search, it could retrieve material more successfully than a human who would have only a catalog to search. This assumption was tested, and under the conditions of the experiment was found to be true.

---

[10] M. J. Ruhl, "Chemical Documents and Their Titles: Human Concept Indexing versus KWIC-Machine Indexing," American Documentation, XV (April, 1964), 136-41.

[11] Don R. Swanson, "Searching Natural Language Text by Computer," Science, CXXXII (October 21, 1960), 1099-1104.

In a later study by Christine Montgomery and Swanson, human indexing operations that take place in an existing system were examined to determine to what extent they could be simulated by machines.[12] Through a detailed examination of a monthly index of medical journals, it was concluded that most of the articles studied could have been indexed by machine on the basis of machine inspection of article titles alone.

In the sense that the investigators of these studies attempted to isolate a human task and establish an automatic model of it, it can be said that the present study closely parallels them. In the area of bibliographic description, and, in particular, formulation of entry, with which the proposed study is associated, simulation has been the subject of some brief discussion by library and information scientists. However, to date exploratory studies and research are lacking. It may be that this lack of investigation stems in part from the belief that practical achievement of this goal is unlikely. The number and variety of cataloging rules and the limitations of existing computers are cited as major obstacles to the automation of these processes.

B. C. Vickery contends that the rules do not fully meet the demand for conciseness that is so important in descriptor files, even though they sum up the principles of accurate and consistent bibliographic description.[13]

---

[12] Christine Montgomery and Don R. Swanson, "Machinelike Indexing by People," American Documentation, XIII (October, 1962), 359-66.

Swanson recognizes the physical limitations of existing hardware, pointing out the general inability of machines to simulate the skills of pattern recognition required in bibliographic cataloging.[14]

C. D. Gull considers the effects of electronic information systems on cataloging rules and procedures and expects the development of new principles, rules, and practices to fit new equipment and new systems.[15] While the new rules may not be suitable for smaller libraries well served by card catalogs, he suggests that the old cataloging rules will need to be made compatible with the new rules.

Finally, participants at the Second Conference on Machine-Readable Catalog Copy discussed the possibility of simulating cataloging rules by machine techniques, but were reminded that the diversity and non-uniformity of title pages would prevent mechanization of these processes.[16]

---

[13] B. C. Vickery, _On Retrieval System Theory_ (London: Butterworths, 1965), pp. 20-21, 125-51.

[14] Don R. Swanson, "Library Goals and the Role of Automation," _Special Libraries_, LIII (October, 1962), 470.

[15] C. D. Gull, "How Will Electronic Information Systems Affect Cataloging Rules?" _Library Resources and Technical Services_, V (Spring, 1961), 135-39; see also his "[Working Paper] No. 17, The Impact of Electronics Upon Cataloging Rules," in _The Report of the International Conference on Cataloguing Principles, Paris, 9th-18th October, 1961_ (London: International Federation of Library Associations, 1963), pp. 281-90.

In fields other than librarianship, the usefulness of simulation for analyzing the performance of human tasks has been amply demonstrated. Since the research is largely exploratory and necessarily narrow in scope, the research results are considered suggestive rather than final. In areas such as machine translation where there has been extensive research, certain findings have been well established.[17] In most fields, experimental work has been too limited to permit a general statement of the findings.

Summary

In summary, the problem set for this study was to examine the bibliographic data and cataloging rules used in the assignment of the main entry, and to construct a formal, theoretical model that would simulate the human process of selecting this entry. Simulation experiments were reported in a large number of fields. Except for indexing, application of the technique to library processes has been minimal. A few studies in indexing were cited for their relevance to this study. In the area of bibliographic description and, in particular, formulation of entry, with which the proposed study is associated, simulation has been the subject of some brief discussion by library and information scientists. The generally conceded complexity of these processes was cited as an obstacle to their automation.

---

[16] Proceedings of the Second Conference on Machine-Readable Catalog Copy, Library of Congress, November 22, 1965 (Washington, D.C.: Library of Congress, 1965), pp. 14-16.

[17] See, Yehoshua Bar-Hillel, Language and Information (Reading, Mass.: Addison-Wesley Publishing Co., 1964), pp. 153-218; Language and Machines: Computers in Translation and Linguistics (Washington, D.C.: National Academy of Sciences, 1966), 124 p.

In fields other than librarianship, the usefulness of simulation for analyzing the performance of human tasks has been amply demonstrated. Since the research is largely exploratory and necessarily narrow in scope, the research results are considered suggestive rather than final.

CHAPTER II

HYPOTHESIS AND GENERAL RESEARCH METHOD

This chapter will discuss the general hypothesis that guided the study and the methods and procedures followed for testing the hypothesis. In the discussion of the latter, there will be a description of the techniques used to simulate human performance of the entry selection process and the methods and procedures used to test their practical application to bibliographic materials.

The Hypothesis

The specific cataloging process selected for simulation by automatic techniques was that which concerns selection of entry or the determination of authorship responsibility. The problem of entry as conceived and formulated in the cataloging codes is specifically a decision problem in that it requires of the cataloger (a) maximum recognition of the presence or non-presence of a particular set of bibliographic conditions, and (b) the need to decide the entry that is valid for these conditions.

In a more formal sense, the process can be likened to an environment specified by a core of bibliographic principles which serve to govern the selection process, by a body of enumerative rules which define the features of the bibliographic conditions to which the governing principles apply, and by an undetermined quantity of bibliographic materials which form the corpus for the application of the rules and principles. Within this environment, the decision or selection process is accomplished through a series of logical, sequential operations which the cataloger performs with each bibliographic item. These operations can be envisioned as involving a sequence, or rather sequences, of questions, answers, and entry choices based on the cataloger's examination of relevant bibliographic data. For instance, to determine the presence of authorship initially, a cataloger would presumably ask, "Is there an author name?" If there were two author names, the typography presumably would be questioned: "Is one author name in larger type?" and so on. In such a procedure, the questions presume two values: positive and negative values or instances of authorship.

The content and structure of these operations suggest that a formal representation of this process is possible, a representation of a kind that a machine could use to execute the operations performed by the cataloger. A typical representation is a directed tree graph or diagram. Tree graphs occur frequently as operands in data processing algorithms, as, for example, in the analysis of a compound statement, the performance of a classification process, or the execution of a repeated selection sort.[1]

---

[1] See, for example, Earl B. Hunt, Janet Marin, and Philip J. Stone, *Experiments in Induction* (New York: Academic Press, 1966), pp. 16-23; David G. Hays, *Introduction to Computational Linguistics* (New York: American Elsevier Publishing Co., 1967), pp. 31-38; G. Salton, "The Manipulation of Trees in Information Retrieval," in *Information Storage and Retrieval System*, Report ISR-1, Section IV (Cambridge, Mass.: Computation Laboratory, Harvard University, 1961), pp. 26-28.

A directed tree graph is characterized much in the way its name implies--by a root, nodes or leaves, and associations between pairs of such nodes expressed as branches or branching tree structures. At the ends of the branches are terminal nodes indicating positive or negative endpoints in a sequential analysis. A directed tree graph admits of a simple graphical representation, as illustrated on the following page (Figure 1). The nodes are depicted as circles, and the branches as directed lines. The structure of the graph is represented entirely by the directed lines.

In selection of entry, a directed tree graph would have as its nodes the relevant bibliographic features of the rules which characterize the governing principles, the two lines of branches extending from the nodes the positive and negative values of the features, and the terminal nodes the decision or entry that is valid for the presence or absence of these features. The nodes and branches would serve to characterize not only the content and sequential operations of the process, but would also act as operands on the bibliographic data presented to the system. When converted to matrix or some equivalent form, they would facilitate machine manipulation. In graphic form, they would facilitate comparable automatic manipulation by humans lacking both professional library training and work experience.

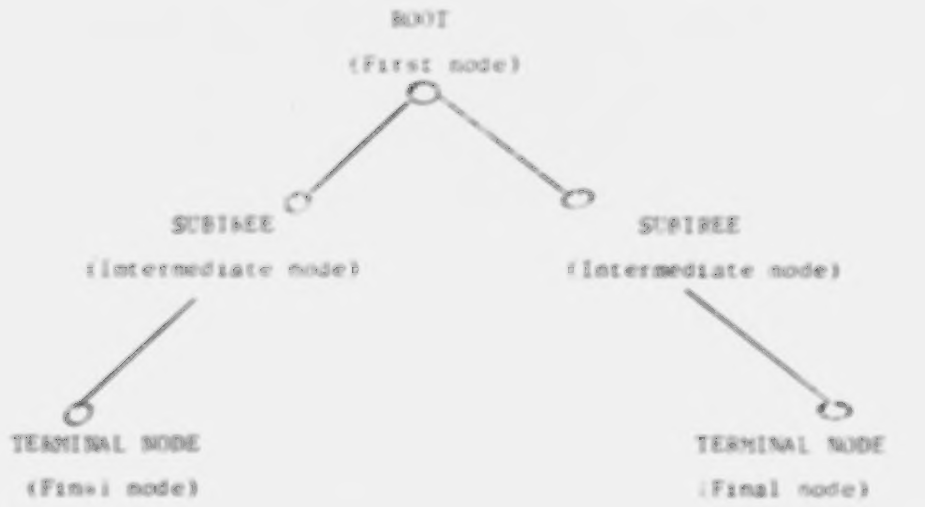

ROOT
(First node)
SUBTREE
(Intermediate node)
SUBTREE
(Intermediate node)
TERMINAL NODE
(Final node)
TERMINAL NODE
(Final node)

Figure 1. Abstract Tree

Briefly defined, an abstract tree can be considered a special type of directed graph consisting of elements (nodes) and associations between pairs of elements (branches). A branch from, say, node i to node j is said to *leave* node i and to *enter* node j. Since each node is entered by at most one branch, a path existing between any two nodes in a tree is unique, and the length of the path is likewise unique. The first node of the tree is called the *root* of the tree and the final node of any branch of the tree is the *terminal* node. Any node of the tree other than the terminal node defines a *subtree* of which it is the root, consisting of itself and all nodes reachable from it. Every subtree of a tree then is itself a rooted tree.[2]

---

[2] Terminology and definitions are adapted from G. Salton, "The Manipulation of Trees in Information Retrieval," in *Information Storage and Retrieval System*, Report ISR-1, Section IV (Cambridge, Mass.: Computation Laboratory, Harvard University, 1961), pp. 26-28.

Such analysis and formulation of the entry process infer that human selection of entry can be simulated by machine or comparable automatic techniques. However, a question arises in connection with the revisions of the cataloging rules used for selection of entry. It would be reasonable to suppose that amenability to automatic simulation would be affected by these revisions. To the extent that the content of the rules in the old and new codes permits equally the abstraction of the basic underlying principle governing the choice of entry and a feature analysis of the relevant bibliographic conditions which define that abstraction, all rules for entry can be said to be amenable. Nonetheless, within the context of this fundamental disposition to simulation, it is recognized that differences in the content, number, design, and structure of the rules in these codes could affect the complexity of the trees, their operational capacity, and ability to make appropriate entry selections. Indeed, it is precisely on the basis of these differences that the new rules for entry contained in the Anglo-American cataloging code are said to be an improvement over former rules in the ALA code, thereby suggesting that the new rules for entry in the Anglo-American code are more amenable to simulation by automatic techniques than those in the ALA code.[3] It is in this connection that investigation was made of the hypothesis that

---

[3] By Anglo-American cataloging code and ALA cataloging code, reference is made specifically to The American Library Association and others, Anglo-American Cataloging Rules, North American Text (Chicago: American Library Association, 1967), 400 p., and The American Library Association, Division of Cataloging and Classification, ALA Cataloging Rules for Author and Title Entries, 2d ed., ed. by Clara Beetle (Chicago: American Library Association, 1949), 265 p. For these codes, the abbreviation AA and ALA, respectively, will be used throughout the thesis.

> A corpus of rules used for selection of entry in the Anglo-American code is more amenable to simulation by automatic techniques than its counterpart corpus of rules in the ALA code, because the corpus in the Anglo American code contains fewer rules; the rules are less complex, less redundant, and enumerate more specifically the bibliographic conditions operative for selection of entry. Because of these factors, the corpus of rules for selection of entry in the Anglo-American code will produce a branching tree diagram which will be less complex and which will generate the correct choice of entry for a greater number of bibliographic materials than will the tree diagram by the ALA corpus of rules.

To facilitate a more accurate interpretation of the hypothesis, explanations are provided for the use of several terms.

A corpus of rules refers to a group of rules selected for formal characterization in the study. The selection of each corpus was based on a number of considerations: comparison of content, sequence and development of the rules, potential characterization, etc.

Simulation of the entry process by automatic techniques refers specifically to the use of those mechanistic devices and procedures which will facilitate either machine or comparable automatic manipulation of cataloging data to produce the proper entry selection. The mechanistic devices used in this study are the branching tree structures already mentioned. The mechanistic procedures are the heuristics developed in the trees. In the sense that entry selection is accomplished objectively and non-intuitively, it is said to be automatic.

With respect to the cataloging codes, rules are considered complex in terms of the number of conditions they specify. These conditions have their verbal statements in the rules as, for example: "Enter collected letters of one writer under the writer" (ALA 6B), where "collected letters" and "one writer" are two conditions of entry. A rule with five conditions, for instance, would be considered a much more complex rule than one with two conditions. In turn, a corpus of rules having a total of fifty conditions would be considerably more complex than one with twenty-five conditions. With respect to the trees, branching structures are considered less complex in terms of the number of nodes and subtrees required for representation of the cataloging rules and principle of entry. Representation of an ALA rule of entry requiring three so-called subtrees would be considerably more complex than, say, the representation of a comparable AA rule requiring only one subtree.

Redundancy in the rules refers to unnecessary repetition or duplication of conditions and specifications of entry, for example, the enumeration of specific rules for specific cases as illustrated in ALA 23A: "enter a dramatization based on a novel...under the playwright...," and its companion, ALA 23B: "Enter a novelization based on a play...under the novelist..."

Bibliographic conditions are considered to be enumerated specifically if they permit objective observation in contrast to subjective, intuitive interpretation. For instance, the condition of "principal responsibility" posited as such is an obscure, intuition-bound notion, subject to the interpretation and philosophical insights of a cataloger. On the other hand, with the specification that principal responsibility is the first-named author on the title page, the name typographically or lexically prominent, etc., explication of the notion is in terms of direct, observable constructs, independent of any special semantic significance or interpretability. In this sense, the condition would be said to be enumerated specifically.

General Research Method

Before discussing the methods developed in this study, certain assumptions and restrictions basic to the whole approach of the study require clarification.

Central to the development and operational capacity of the simulation model used in the study is the assumption that a device exists for reading directly from a document. More specifically, it was assumed that this device accepts the title pages of English language monographs, and simulates the human reading of the printed characters and symbols on that page, namely, reading sequentially from left to right, top to bottom. Since the title and author statements constituted the cataloging data on which the selection of entry was based, the device would need to be restricted to reading these statements only.

Central to the construction and application of the simulation model was the assumption that rules could be devised which would enable this reading device to recognize or detect the presence or absence, as the case may be, of authorship on the title page. This assumption was tested in the study.

In view of the various dimensions of the problem, the study was addressed primarily to conceptual problems at the base of the entry process, and the construction of a device that would incorporate the design and structure of this process. While the investigation of the process did not preclude the importance of those issues that relate to technical and economic feasibility, these issues were not considered.

In placing these constraints on the study, certain problems relating to these issues did not arise. For instance, a fundamental problem would be the availability of an electronic device with the sophistication and flexibility of operation needed to perform the simulation process; similarly, the "building" or programming of this device would be necessary to execute the operations determined by the simulation model. In short, problems of hardware and cost were not within the scope of the investigation although their relevancy was not denied.

With respect to the simulation model a severe constraint was placed on its formulation, namely, that a formal characterization of the entry process was not simultaneously a model of the actual performance of that process by human catalogers. For example, the fact that catalogers are able to comprehend an unlimited variety of novel title page data and can judge and determine their usefulness in formulating an entry choice indicates immediately that their capacities are quite different from an automaton that performs operations on the basis of printed characters and symbols presented to it. Consequently, there was no claim that the simulation model developed in the study was a model of the intellectual or intuitive processes of the human cataloger.

The general research method used to test the hypothesis involved studying and analyzing the principles and rules for author and title entry, and the bibliographic data used in the formulation of these entries in the AA and ALA codes. In the case of both codes, this was done by writing down descriptive statements in formal terms which were derived from the rules, and by specifying those elements of bibliographic data that were used by the rules to determine the entry choices.

On the basis of this analysis two sets of rules were formulated for each code. The function of the first set of rules was to "recognize" or detect the presence of authorship on the title pages of English language monographs. These rules were called in terms of their function "recognition rules". The data for these rules were certain lexical units extracted from the title and author statements that were found to accommodate entry selection in the two codes. The function of the second set of rules was the determination of author responsibility. These rules were called in terms of their nature and function "experimental selection rules". The data for these rules were the specifications for selecting entry contained in the verbal principles and rules of both codes. A formal, graphic representation of these rules was derived by mapping these specifications into directed tree graph structures, where the nodes of the structures contain the specifications, and the endpoints of the structures are positive and negative selection choices representing the abstraction of entry principles.

A component of the representation was the heuristic procedures, formulated and applied to the representation, that enabled it to sample the bibliographic data produced under the application of the recognition rules in order to derive the proper selection choices. In the study, these procedures involved the use of a binary (YES - NO) decision process, commonly used to represent machine computations.

Taken together, then, the recognition rules, the selection rules, and the heuristic component comprised the simulation model for the human ordering of the entry process.

The degree of amenability of the rules of entry in both codes was determined in terms of those variables mentioned: number, complexity, and redundancy of code rules, and the specific enumeration of the bibliographic conditions contained in those rules. This degree of amenability was exhibited in turn in the complexity of the models, and in the ability of the models to accept the bibliographic data on which entry selection was based to generate the correct choice of entry.

In the case of the latter, two tests of adequacy were conducted. The first test, which was performed by the investigator involved the application of each model to a random sample of Library of Congress proof slips. Given certain basic assumptions with respect to the arrangement of the bibliographic data, and relative assurance as to the rules and cataloging code used for the printed entries, the ability of each model to generate correct entry choices was determined. The adequacy of the models to accomplish entry selection automatically, that is, objectively and non-intuitively, was demonstrated in a second test involving an experiment, in which the simulation models were applied to a specially selected problem set of English language monographs. To carry out the experiment, a group of six persons, totally lacking in professional library training and experience, was formed to function as an automaton in the entry selection process. On the basis of only an introduction to the model rules and tree structures, they served to demonstrate that selection of entry could be accomplished in an automatic way.

In an attempt to determine the effectiveness of the technique insofar as it assists professionally trained catalogers to formulate entry choices automatically, a second group of six catalogers was formed. This group was assigned the same task as the first group, using the same problem set of English language monographs.

In the case of each group, the rules and trees were marked for each monograph to show the heuristics used to derive entry selection. In terms of performance measures, this permitted observation and analysis of the interactions that took place between the designated input or cataloging statements, the sequence and structure of particular problems, and the way in which these related to mechanistic efficiency. It also permitted comparison of the two models and of the two groups performing the experiment. As an index of accuracy, Library of Congress printed cards were used for the problem set of English language monographs.

Summary

To summarize, entry selection as conceived and formulated in the cataloging codes can be described as a decision problem, in that it requires on the part of the cataloger the recognition of the presence or non-presence of a particular set of bibliographic conditions, and the need to decide the entry that is valid for these conditions. This decision process was envisioned as a series of questions, answers, and entry choices based on the cataloger's examination of relevant bibliographic data. The content and structure of these operations suggest that human selection of entry can be accomplished in an automatic way using an appropriate automatic technique.

Directed tree graphs constitute a machine technique commonly used to simulate human learning and decision tasks. Rules of entry in both the AA and ALA cataloging codes were considered amenable to simulation by this technique. However, it was recognized that differences in the content, number, and structure of the rules could affect the complexity, operational capacity, and ability of the trees to make appropriate entry selections. On this basis, it was hypothesized in general that a corpus of rules for selection of entry in the AA code would be more amenable to simulation by automatic techniques than its counterpart corpus of rules in the ALA code. Definitions of certain terms used in the hypothesis were given, as well as a statement of certain assumptions and constraints basic to the whole approach of the study. The general research method used to test the hypothesis was discussed, and procedures for constructing the simulation models, and for testing their degree of amenability and adequacy were outlined.

CHAPTER III

RULES OF ENTRY: SELECTION AND AMENABILITY

This chapter will explain in detail the research method used to test the hypothesis. In particular, there will be a discussion of the rules of entry in the AA and ALA codes that were selected for characterization in the study, the basis for their selection, and the measurement of the degree of their amenability to simulation by automatic techniques. The first section will discuss the procedures used to select a corpus of rules from the two cataloging codes and the considerations that influenced the final selection of each corpus. The second section will describe the procedures employed to measure the degree of amenability of each corpus of rules to simulation by automatic techniques, followed by a statement of the findings resulting from these procedures.

Selection of Rules

Rules of entry selected for characterization in the study were from Chapter One in the AA and ALA codes. In view of the number of these rules, thirty-two in the AA and thirty-five in the ALA codes, not including their alphabetical or numerical parts and sub-parts, it was necessary to select a corpus of rules from each code. The basis for the selection of this corpus involved a number of considerations. A representative corpus from each code would presumably reflect frequency of rule use and entry type, and perhaps give some attention to rule order and development. Of particular importance is that the corpus from both codes would permit comparison of content, and at the same time provide its formalized model of entry selection with sufficient descriptive devices for the model to demonstrate both its ability and potential to select appropriate main entry choices.

Regrettably, there are no published statistical data on the frequency of use of particular rules of entry that would provide independent justification for the selection of each corpus. An inquiry directed to the Library of Congress revealed that the Library makes no attempt to keep these statistics.[1] Similar inquiries to several large academic and public libraries produced the same response. In a statistical study in which the cataloging data on Library of Congress cards were examined in connection with determining fields of information for a machine-readable format, the types of main entry and the frequency of their occurrence in two random samples were reported.[2] While this information was of interest, it could not be used to form a reliable judgment of rule use. In the absence of these data, the selection of each corpus was initially determined on the basis of the distribution of rule and entry type observed in a random sample of Library of Congress English-language proof slips. For each corpus, two weekly shipments of proof slips were used for the examination. Since the shipments were received at the University of Illinois library school under a blanket order placed with the Library of Congress, the contents were not biased by institutional or personal selection. There were two reasons for using two shipments for each corpus: first, in terms of sample size, the five to six thousand proof slips that two shipments were likely to yield were considered to provide a sufficiently large base on which to initiate the selection of the rules for characterization, and second, in terms of bias, the use of two shipments lessened the possibility of an accidental preponderance of a particular entry type, or the use of a particular entry rule.

---

[1] Letter from John C. Rather, Specialist in Technical Processes Research, Library of Congress, October 8, 1969.

[2] Henrietta D. Avram, Kay D. Guiles, and Guthrie T. Meade, "Fields of Information on Library of Congress Catalog cards: Analysis of a Random Sample, 1950-1964," The Library Quarterly, XXXVII (April, 1967), 188.

With respect to the cataloging codes used for the entries in each set of shipments, a measure of assurance was the date, March 20, 1967, taken by the Library of Congress for its application of the AA code.[3] Cards issued before that date could be presumed to have been established under ALA rules, while cards issued after that date could be presumed to have been established under AA rules. With respect to the latter, two problems were encountered. First, since the printing of Library of Congress cards does not necessarily occur in the same order as the descriptive cataloging, a number of cards issued after that date could have been cataloged under ALA rules, a situation that the Library of Congress forewarned catalogers about in its bulletin.[4] Second, the policy of "superimposition" adopted by the Library of Congress limits the application of the AA rules of entry only to works which are new to the Library.[5] Since the Library does not indicate the entries affected by this policy, all cards must be examined for possible conflicts in entry choices. Consequently, the following precautions were exercised with respect to the selection of the sample. shipments of proof slips received at the library school prior to 1966 were considered to be catalogued unquestionably under ALA rules of entry, while shipments received in 1968 and later were considered partial assurance that the AA rules had been used to assign the main entries. The two shipments used to determine rule use in the ALA code were received in the months of October and December, 1965. Those used to determine rule use in the AA code were received in the months of June and August, 1969.

---

[3] U. S. Library of Congress, Processing Department, Cataloging Service, Bulletin No. 80 (April, 1967), 1.

Examination of the entries was first made by the investigator, followed by an independent examination, which was performed by a cataloger with experience and knowledge of the rules in the two codes. In the case of each entry, the code, rule or rules, and entry type were designated. Where an agreement was not reached, the entries in question were examined by an additional independent source, and whenever possible, the titles were secured for examination. The results of this examination are given in the following tables. Tables 1 and 2 show the frequency and rank of difference in the use of main entry rules in the AA and ALA codes respectively. As expected, the general rules in both codes ranked the highest in use. In line with the brevity of the code, it could be expected that there would be fewer AA rules used. As the totals show, for the two samples of comparable size, 63 AA rules were applied in contrast to 134 ALA rules.

---

[4] Ibid.

[5] Ibid., Bulletin No. 79 (January, 1967), 1.

Table 3 shows the frequency of main entry type for both samples. In general, the distribution was predictable. In both codes, personal author monopolized the main entry, followed by corporate author, title and uniform headings (Bible, anonymous classics, etc.). Less predictable, perhaps, was the larger percentage of title main entries in the AA sample, 7 percent in contrast to 5 percent in the ALA sample. It is tempting to attribute this difference to particular rules in the AA code. In the shipments for each sample, for instance, the distribution of main entries was closely related, thus lessening the introduction of bias. Table 4 combines a frequency count of the main entries in these shipments. In the statistical study done at the Library of Congress, which was cited earlier, the two samples examined for their frequency of main entry type involved application of the ALA rules. A comparison of the distribution of those samples and the shipments comprising the ALA sample used in this study showed them to be closely related.[6] However, reference to Tables 1 and 2,

---

[6] Avram, Guiles, and Meade, op. cit., p. 188.

Table 1

Frequency of Use of AA Rules of Entry

AA Rule	Library of Congress Proof Slips		Rank
	Number	Percentage	
1A	3444	56.89	1
2A	110	1.82	14
2B	3	.05	28
2C	1	.017	30
3A	73	1.21	17
3B1	506	8.36	2
3B2	48	.79	18
4A	222	3.67	5
5A	189	3.12	8
5B	3	.05	28
6A	128	2.11	11
6B1	105	1.73	15
6B2	113	1.87	13
6C	1	.017	30
6D1	7	.111	25
7A	15	.25	23
8A	7	.111	25
8B	305	5.04	3
8D1	2	.034	29
8D2	1	.017	30
8E1	5	.08	27
8E2	8	.13	24
9A	3	.05	28
9B	3	.05	28
11A	3	.05	28
11B	2	.034	29
13A	2	.034	29
13B	2	.034	29
13C	1	.017	30
14A	120	1.98	12
14B	1	.017	30
15A	199	3.29	7
16A	3	.05	28
17A1	219	3.62	6
17A2	31	.51	20
17B	144	2.38	10
17C1a	1	.017	30
17C1b	1	.017	30
18A	247	4.08	4

Table 1 (continued)

AA Rule	Library of Congress Proof Slips Number	Percentage	Rank
18B	176	2.91	9
19A1	1	.017	30
19A2	2	.034	29
19B	37	.61	19
19B2	2	.034	29
20A	21	.35	22
20B1	1	.017	30
20B2	2	.034	29
20C	82	1.35	16
21A	1	.017	30
21B	1	.017	30
22A1	6	.10	26
22A2	1	.017	30
23B	1	.017	30
23D	3	.05	28
23E	1	.017	30
24B	3	.05	28
25A1	23	.37	21
25A2a	1	.017	30
25B1	1	.017	30
25F1	1	.017	30
26B	1	.017	30
27A	2	.034	29
29A1	2	.034	29

Total Rules of Entry Used	63
Sample of LC Proof Slips	6054

Table 2

Frequency of Use of ALA Rules of Entry

	Library of Congress Proof Slips		
ALA Rule	Number	Percentage	Rank
2	3487	58.43	1
3A	593	9.94	2
3B	71	1.19	13
3D	2	.03	43
3E	13	.22	32
4A(1)	86	1.44	10
4A(2)	23	.39	25
4B(1)	5	.08	40
4B(2)	1	.017	44
5A(1)	261	4.37	4
5A(1a)	14	.23	31
5A(1c)	12	.20	33
5A(2)	5	.08	40
5B	4	.07	41
5C(1)	120	2.01	8
5C(3)	2	.03	43
5C(4)	22	.37	26
5D	3	.05	42
5E(1)	1	.017	44
5E(2)	4	.07	41
6A	7	.17	38
6B	12	.20	33
6B(1)	11	.18	34
6B(2)	5	.08	40
6C	2	.03	43
10A	59	.99	14
10B	25	.42	23
11	2	.03	43
12A	84	1.41	11
12A(1)	2	.03	43
12A(2)	9	.15	36
12A(3)	2	.03	43
12A(5)	1	.017	44
12A(7)	3	.05	42
12A(8)	4	.07	41
12A(10)	1	.017	44
12A(11)	2	.03	43
12B	9	.15	36
12C	2	.03	43

Table 2 (continued)

ALA Rule	Library of Congress Proof Slips Number	Percentage	Rank
12E	49	.82	17
13A	1	.017	44
13C	5	.08	40
14A	2	.03	43
14C	1	.017	44
16	1	.017	44
17A	24	.40	24
19A	24	.40	24
19D	5	.08	40
19E	309	5.18	3
20	19	.33	27
21A	238	3.99	5
21B	1	.017	44
22A	2	.03	42
22B	15	.25	30
22C	86	1.44	10
23A	2	.03	43
23B	5	.08	40
24	1	.017	44
25A	8	.13	37
25C	6	.10	39
26	4	.07	41
27B	1	.017	44
27F	1	.017	44
28	1	.017	44
29A	2	.03	43
30A	28	.47	22
30A(1)	8	.13	37
31B	2	.03	43
31B(2)	1	.017	44
32GR	10	.17	35
32B	1	.017	44
33GR	14	.23	31
33A	1	.017	44
33A(1)	1	.017	44
33E	1	.017	44
34	3	.05	42
35B	2	.03	43
35C(1)	1	.017	44
71	169	2.83	6
72	34	.57	20
72A	151	2.52	7
72A(1)	32	.54	21
72B	1	.017	44

Table 2 (continued)

ALA Rule	Library of Congress Proof Slips		Rank
	Number	Percentage	
73A	2	.03	43
73B	1	.017	44
73C	1	.017	44
74	50	.84	16
75	12	.20	33
75A	83	1.39	12
75B(1)	10	.17	35
75B(2)	50	.84	16
75C	112	1.87	9
76B	12	.20	33
81A	19	.33	27
84	17	.29	29
84B	5	.08	40
84C	4	.07	41
84D	3	.05	42
84E	2	.03	43
85	3	.05	42
87	1	.017	44
88A	14	.23	31
88D	2	.03	43
89H	1	.017	44
90D	1	.017	44
91GR	34	.57	20
91B(1)	5	.08	40
91B(2)	10	.17	35
92GR	42	.70	19
92A	22	.37	26
93E	1	.017	44
96	2	.03	43
97	10	.17	35
101A	1	.017	44
102	7	.12	38
104	24	.40	24
116F(4)	1	.017	44
119A	2	.03	43
120A	1	.017	44
120C(1)	1	.017	44
120D	1	.017	44
135A	55	.92	15
135B	4	.07	41
135C	50	.84	16

Table 2 (continued)

ALA Rule	Library of Congress Proof Slips		Rank
	Number	Percentage	
136A	1	.017	44
136B	2	.03	43
138A	1	.017	44
139A	6	.10	39
139B	18	.31	28
143	1	.017	44
144	43	.72	18
144A	24	.40	24
144D	1	.017	44
145	4	.07	41

Total Rules of Entry Used	134
Sample of LC Proof Slips	5968

Table 3

Frequency of Main Entry Type

	Author								
	Personal	Percentage	Corporate	Percentage	Title	Percentage	Uniform Heading	Percentage	Total
AA	4459	73.65	1172	19.36	423	6.99	0	0.00	6054
ALA	4579	76.55	1112	18.63	273	4.57	4	0.06	5968

Table 4

Frequency of Main Entry Type for Shipments Comprising AA and ALA Samples

	Author								
	Personal	Percentage	Corporate	Percentage	Title	Percentage	Uniform Headings	Percentage	Total
AA									
Ship 1	2243	73.37	599	19.59	215	7.04	0	0.00	3057
Ship 2	2216	73.93	573	19.18	208	6.94	0	0.00	2997
ALA									
Ship 1	2242	75.00	567	19.15	150	5.06	1	0.03	2960
Ship 2	2337	77.69	545	18.11	123	4.08	3	0.09	3008

shows a sharp contrast in the samples with respect to the number of proof slips utilizing the rules in the two codes that assign predominantly title main entries. For example, 114 proof slips utilized AA rules 2A, 2B, and 2C, while only 13 proof slips utilized the corresponding ALA rules 31B, 31B(2), and 32. It would appear then that the difference in the assignment of this main entry type might be more attributable to the composition of the samples than to the construct of the two codes.

An important consideration in the selection of each corpus of rules was precisely that their content permitted comparison. Before drawing each corpus from the data then, it was necessary to undertake the identification of parallel rules in the two codes. This task was facilitated by the availability of two published compilations, which were used for reference in the compilation presented in Table 5.[7] The data for the compilation were the rules used in the two samples, and listed in Tables 1 and 2. Matching was accomplished by taking the AA rules of entry listed in Table 1 and recording, where given, the corresponding ALA rules of entry in Table 2. A relatively small number, approximately 60 percent, of the ALA rules were correlated with AA rules. This can be attributed largely to the overlapping of

---

[7] F. Bernice Field, "Anglo-American Cataloging Rules Correlated with A.L.A. Cataloging Rules," in New Rules for an Old Game, ed. by Thelma E. Allen and Daryl L. Dickman (Vancouver: University of British Columbia, 1967), pp. 137-59; Robert B. Slocum and Lois Hacker, Sample Cataloging Forms (2d ed.; Metuchen, New Jersey: Scarecrow Press, 1968), pp. 177-79.

Table 5

Correlation of AA Rules of Entry with ALA Rules of Entry From Tables 1 and 2

AA Rule	IC Proof Slips	Percentage	Rank	ALA Rule	IC Proof Slips	Percentage	Rank
1A	3444	56.89	1	2	3487	54.43	1
				5C(3)	2	.03	43
				16	1	.017	44
				22C	86	1.44	10
				32GR	10	.17	35
				71	169	2.83	6
				136B	2	.03	43
2A	110	1.82	14	31B	2	.03	43
				32GR	10	.17	35
2B	3	.05	28	31B	2	.03	43
				31B(2)	1	.017	44
2C	1	.017	30	32GR	10	.17	35
				32B	1	.017	44
3A	73	1.21	17	3A	593	9.94	2
				3B	71	1.19	13
				4A(1)	86	1.44	10
				12A(3)	2	.03	43
3B1	506	8.36	2	4A(2)	23	.39	25
				4B(1)	5	.08	40
				6A	7	.17	38
				6B(1)	11	.18	34
3B2	48	.79	18	4A(2)	23	.39	25
4A	222	3.67	5	5A(1)	261	4.37	4
				5A(1c)	12	.20	33
				5B	4	.07	41
				5E(2)	4	.07	41

Table 5 (continued)

AA Rule	IC Proof Slips	Percentage	Rank	ALA Rule	IC Proof Slips	Percentage	Rank
5A	189	3.12	8	5A(1)	261	4.37	4
				5A(1a)	14	.23	31
				6C	2	.03	43
				14C	1	.017	44
				21B	1	.017	44
5B	3	.05	28	5A(2)	5	.08	40
6A	128	2.11	11	5C(1)	120	2.01	8
				5D	3	.05	42
				5E(1)	1	.017	44
6B1	105	1.73	15	5C(1)	120	2.01	8
				5D	3	.05	42
				5E(1)	1	.017	44
6B2	113	1.87	13	71	169	2.83	6
6C	1	.017	30	2	3487	58.43	1
6D1	7	.111	25	5C(1)	120	2.01	8
				5C(3)	2	.03	43
				5C(4)	22	.37	26
				5D	3	.05	42
				5E(1)	1	.017	44
7A	15	.25	23	22B	15	.25	30
				23A	2	.03	43
				23B	5	.08	40
				24	1	.017	44
8A	7	.111	25	19D	5	.08	40
8B	305	5.04	3	19E	309	5.18	3
8D1	2	.034	29	19A	24	.40	24
8D2	1	.017	30	----	---	--	--
8E1	5	.08	27	136A	1	.017	44
				136B	2	.03	43

Table 5 (continued)

AA Rule	IC Proof Slips	Percentage	Rank	ALA Rule	IC Proof Slips	Percentage	Rank
8E2	8	.13	24	136A	1	.017	44
				136B	2	.03	43
9A	3	.05	28	6B(2)	5	.08	40
9B	3	.05	28	6B(2)	5	.08	40
11A	3	.05	28	29A	2	.03	43
11B	2	.034	29	29A	2	.03	43
13A	2	.034	29	4B(1)	5	.08	40
				4B(2)	1	.017	44
13B	2	.034	29	4B(1)	5	.08	40
				4B(2)	1	.017	44
13C	1	.017	30	11	2	.03	43
14A	120	1.98	12	20	19	.33	27
				22A	2	.03	42
14B	1	.017	30	20	19	.33	27
15A	199	3.29	7	21A	238	3.99	5
				22B	15	.25	30
16A	3	.05	28	3E	13	.22	32
17A1	219	3.62	6	71	169	2.83	6
				75B(1)	10	.17	35
				135A	55	.95	15
17A2	31	.51	20	71	169	2.83	6
17B	144	2.38	10	5B	4	.07	41
				75B(2)	50	.84	16
				75C	112	1.87	9
17C1a	1	.017	30	73A	2	.03	43
				73B	1	.017	44
				73C	1	.017	44

Table 5 (continued)

AA Rule	IC Proof Slips	Percentage	Rank	ALA Rule	IC Proof Slips	Percentage	Rank
17C1b	1	.017	30	73A	2	.03	43
				73B	1	.017	44
				73C	1	.017	44
18A	247	4.08	4	75A	83	1.39	12
18B	176	2.19	9	75B	---	---	--
19A1	1	.017	30	27F	1	.017	44
19A2	2	.034	29	27F	1	.017	44
19B	37	.61	19	3B	71	1.19	13
				5C(3)	2	.03	43
				5C(4)	2	.37	26
				14A	2	.03	43
				14C	1	.017	44
				25A	8	.13	37
				25C	6	.10	39
				26	4	.07	41
				27F	1	.017	44
				28	1	.017	44
19B2	2	.034	29	12A	84	1.41	11
				12A(2)	9	.15	36
20A	21	.35	22	84	17	.29	29
				84B	5	.08	40
				84E	2	.03	43
20B1	1	.017	30	84	17	.29	29
				84B	5	.08	40

Table 5 (continued)

AA Rule	IC Proof Slips	Percentage	Rank	ALA Rule	IC Proof Slips	Percentage	Rank
20C	82	1.35	16	72A	151	2.52	7
				74	50	.84	16
				84	17	.29	29
21A	1	.017	30	72	34	.57	20
21B	1	.017	30	84	17	.29	29
22A1	6	.10	26	85	3	.05	42
				87	1	.017	44
22A2	1	.017	30	85	3	.05	42
				87	1	.017	44
23B	1	.017	30	---	---	---	--
23D	3	.05	28	---	---	---	--
23E	1	.017	30	---	---	---	--
24B	3	.05	28	84C	4	.07	41
				84D	3	.05	42
25A1	23	.37	21	88A	14	.23	31
25A2	1	.017	30	88A	14	.23	31
25B1	1	.017	30	---	---	---	--
25F1	1	.017	30	---	---	---	--
26B	1	.017	30	89H	1	.017	44
27A	2	.034	29	34	3	.05	42
—				35B	2	.03	43
				35C(1)	1	.017	44
29A1	2	.034	29	116F(4)	1	.017	44

AA Rules of Entry 63
Sample of IC Proof Slips 6054
ALA Rules of Entry 81
Sample of IC Proof Slips 5968

the rules for form of heading with those for choice of entry in the ALA code. As such, many of these rules, especially those for corporate bodies, would be correlated with rules for headings in the AA code.

Differences in the character and construct of the two codes make correlation of the rules difficult. The lack of any stated principles of entry in the ALA code, and the emphasis of that code on specific rules for specific types of publications and classes of authors, are in sharp in contrast to the AA code. Based on a set of stated principles and clearly enumerated objectives, the AA code approaches cataloging in terms of a rational analysis of variant bibliographic conditions which characterize the materials of a library, and provides a core of general rules and a few specific rules to ensure consistent treatment of these conditions.[8] In a few instances, the scope and wording of an AA rule are identical to that of an ALA rule. In most cases, correlation of the two sets of rules can be done only on the basis of determining similarities in points of content, although even in this procedure there are many instances where the content of a rule may be more inclusive or limited than that to which it is related. With few exceptions, the compilation presented here is in agreement with the correlation charts drawn up by the Library of Congress, which were the primary source for the two published compilations cited above

---

[8] See: Seymour Lubetzsky, "Catalog Code Revision, 1964," Library Journal, LXXXIX (December 15, 1964), 4863-65; F. Bernice Field, "The Rules for Entry and Heading," in New Rules for an Old Game, ed. by Thelma E. Allen and Daryl L. Dickman (Vancouver: University of British Columbia, 1967), pp. 33-43; James A. Tait, Authors and Titles (Hamden, Conn.: Archon Books, 1969), pp. 121-39.

Additional data from Tables 1 and 2 incorporated in the compilation suggest points of comparison in the composition of the two samples Although a great deal of significance cannot be attached to these figures, it is interesting to note the overall similarity in the composition of the two random samples. This is evidenced by the numbers of proof slips affected by sets of comparable rules and also in the ranking of rule use. An obvious example of the is AA rule 8B and its ALA counterpart, 19E. In both the number of proof slips and rank of usage, the two rules were clearly correlated. Similarly in the case of AA 1A and ALA2, AA 5B and ALA 5A(2), and most of the general rules in both codes, close correlation is noticeable.

A final and important consideration in the selection of each corpus was that the content of the corpus would be sufficiently descriptive for the formalized model to demonstrate both its ability and potential to produce appropriate main entry choices. This consideration had the effect of limiting the size of the corpus while at the same time expanding the base for the inclusion of a variety of rules. For example, consider the following rules:

(1) Enter correspondence between two or more persons under the first one mentioned on the title page ... (ALA 6)

(2) Enter under corporate body, a work that is by its nature necessarily the expression of the corporate thought or activity of the body ... (AA 17A1)

From the viewpoint of formal characterization, the content of the first rule was considered to be sufficiently descriptive, that is to say, it is explicit in the specification of exact data (correspondence, two persons, first one mentioned, etc.) to permit objective "mechanical" selection of an appropriate main entry type. In just what way the data would facilitate a device to perform the selection, and in what sense selection would be mechanical are taken up in detail in Chapter IV. For the purposes of the present discussion the explanationsare intended to be illustrative only.

In the same context of formal characterization, the content of the second rule was considered insufficiently descriptive; that is to say, it is inexplicit in the specification of the data (corporate body, expression, thought, etc.). While the rule may be clear to a cataloger, it is formulated in highly abstract terms. It is only after answers to such questions as "What is a corporate body?", "What constitutes an expression of corporate thought?", etc. are specified in detail that the descriptive data are made exact and the content of the rule explicit.

As a result of this examination of the rules, a number of entry rules from both codes were excluded. These encompassed the rules as well as works of corporate authorship, serial, legal, and religious publications, as well as certain parts of the general rules listed in the compilation in Table 5. Due to their physical character, non-book materials were also excluded. A brief look at the figures for sample use of these rules in the table reveals that with the exception of one or two rules for corporate bodies and serials, the excluded rules, on a whole, ranked low in usage.

A seemingly contradictory result of this examination was that in spite of its limiting character, it expanded the base for the inclusion of a broad choice of rules in each corpus. For instance, it was noted that in the case of each code, a number of rules shared certain descriptive data. One clear example is ALA 22A:

> Enter an abridgment, epitome, or outline under the same heading as the original work ...

and 22B:

> Enter an adaption, a paraphrase, or a free translation, as a general rule, under the same heading as the original work ...

where both rules share the feature of "original", that is, entry under the original author. By using the same descriptive data, a number of rules were in effect collapsed into one rule, which made it possible to consider the formalization of a variety of rules for each corpus. It also served to demonstrate the potential of the formalized models to make appropriate entry choices for a broad range of materials. A more detailed explanation of the rules and the kinds of materials requiring their application will be provided in Chapter IV.

On the basis of these considerations, a corpus of nineteen AA rules of entry and thirty-six comparable ALA rules of entry were selected for characterization in the study. These rules and the data on their frequency of use in the two samples are listed in Table 6. For reference to the rules themselves, see Appendix A, pp. 311 - 25. In the case of each sample, these included the three most frequently used rules for personal authorship, which accounted for 58 to 73 per cent of sample rule usage. Finally, the inclusion of other rules was based on a number of considerations: comparison of content, sequence and development in the code, integration into the corpus, and formalization of the descriptive data.

Measurement of Amenability

It was proposed that there were certain specific factors associated with the content and structure of each corpus that determined the degree to which it would be amenable to formal characterization. In particular, it was specified that variables of size, complexity, redundancy, and specificity affect the formal characterization of a corpus, determining its complexity, operational capacity, and ability to produce appropriate main entry choices in an automatic way. In this connection, it was hypothesized that the corpus of AA rules of entry would be more amenable to formal characterization than its counterpart corpus of ALA rules, precisely because there are fewer rules in the corpus; the rules are less complex, less redundant, and enumerate more specifically the bibliographic conditions operative for selection of entry.

Table 6

Corpus of AA Rules of Entry and Comparable ALA Rules of Entry
Selected for Characterization

AA Rule	LC Proof Slips	Percentage	Rank	ALA Rule	LC Proof Slips	Percentage	Rank
1A	34444	56.89	1	2	3487	58.43	1
				16	1	.017	44
				22C	86	1.44	10
2A	110	1.82	14	32GR	10	.17	35
2C	1	.017	30	32GR	10	.17	35
				32B	1	.017	44
3A	73	1.21	17	3A	593	9.94	2
				3B	71	1.19	13
				4A(1)	86	1.44	10
3B1	506	8.36	2	4A(2)	23	.39	25
				4B(1)	5	.08	40
				6A	7	.17	38
				6B(1)	11	.18	34
3B2	48	.79	18	4A(2)	23	.39	25
4A	222	3.67	5	5A(1)	261	4.37	4
				5A(1c)	12	.20	33
				5B	4	.07	41
				5B(2)	4	.07	41
5A	189	3.12	8	5A(1)	261	4.37	4
				5A(1a)	14	.23	31
				6C	2	.03	43
5B	3	.05	28	5A(2)	5	.08	40
7A	15	.25	23	22B	15	.25	30
				23A	2	.03	43
				23B	5	.08	40
				24	1	.017	44

Table 6 (continued)

AA Rule	LC Proof Slips	Percentage	Rank	ALA Rule	LC Proof Slips	Percentage	Rank
8A	7	.111	25	19D	5	.08	40
8B	305	5.04	3	19E	309	5.18	3
9A	3	.05	28	6B(2)	5	.08	40
9B	3	.05	28	6B(2)	5	.08	40
14A	120	1.98	12	20	19	.33	27
				22A	2	.03	42
14B	1	.017	30	20	19	.33	27
15A	199	3.29	7	21A	238	3.99	5
16A	3	.05	28	3E	13	.22	32
19B	37	.61	19	5C(4)	22	.37	26
				14A	2	.03	43
				25A	8	.13	37
				25C	6	.10	39
				26	4	.07	41
				28	1	.017	44

AA Rules of Entry	19
Comparable ALA Rules of Entry	36

At this point, it is possible to state a little more precisely what is required in the way of a formal characterization of the rules, although procedures for the actual construction of the characterization will be taken up in Chapter IV.

A fundamental aim of the characterization of each corpus was the construction of a formalized model of entry selection that (a) would specify explicitly the appropriate main entry choices for the bibliographic data presented to it, and (b) at the same time would meet certain conditions of adequacy and simplicity which would provide justification for the choice of one model of entry selection over the other.

A prerequisite of the characterization was that it be formal. In the sense that it uses certain descriptive devices (letters, letters combined into word units, numbers, and binary features ±) of a language and not the meanings or semantic connotations of a language, the model for a corpus can be said to be formal. Thus, the statement that a corporate body is any organization or group of persons that is identified by a name is not a formal statement. It might be offered or appended as an explanation, but could not be part of the characterization. In order to be incorporated into the proposed characterization, it would have to be translated into formal terms, namely, through the use of the descriptive devices already mentioned.

A further requirement of the characterization was that it be explicit. In the sense that it provides a series of "mechanical", counter-subjective, non-intuitive steps to accomplish the selection of the main entry, the model for a corpus can be said to be explicit. While a corpus of verbal rules will depend heavily on the judgment, accumulated experience, and intuition of the user to interpret its statements, a characterization of the type proposed can do no such thing.

Finally, it was a requirement of the characterization that it be simple. To the extent that the characterization fits or accounts for the data with an economy of descriptive devices and "mechanical" steps, the model for a corpus can be said to be simple. Thus, from the viewpoint of justifying a choice of models, given two that account for equivalent data, the one that is simpler (more economical, in the sense of containing more general statements and fewer special ones) is naturally preferred. However, it is important to note that simplicity in one respect can carry with it complexity in another. The criterion is meaningful only in the context of the whole formalization. There will be more to say about this and the other requirements later. At this point, it is essential only that the aims of characterization are set significantly for the testing of the hypothesis that follows.[9]

The testing of the hypothesis called for measuring the amenability of each corpus in terms of the four criterion variables mentioned. size, complexity, redundancy, and specificity. Central to the measurement was the formulation of a rigorous, precisely stated definition of each criterion variable that would facilitate testing on an objective basis. The difficulty in developing the formulation was recognized at the outset. A crucial difficulty was the lack of any naturally suggested scales for at least three of the variables involved. For instance, it is not at all clear a priori how complexity, redundancy, and specificity can be measured. Although such measures are not unthinkable, it was recognized that none would be a priori more justifiable than others. Thus, while it was possible to suggest operational definitions of the criterion (which would permit quantitative measurement), their justification, in fact, would have to be made a posteriori, that is, in view of the adequacy of the characterization. For instance, a corpus of rules could be determined complex in terms of certain quantitative measures specified by a definition, but unless these were observed in the characterization and found to affect it in some determinable way (ratio of nodes, branching in the trees, etc.), then the operational definition posited for that variable, or even the variable itself, would not be justified as a relevant factor of amenability.

---

[9] From the viewpoint of model construction, these requirements are generally accepted for ideal characterization. See, for example: Anatol Rapoport, "Mathematical Models of Social Interaction", in Handbook of Mathematical Psychology, ed. by R. Duncan Luce, Robert R. Bush, and Eugene Galanter (New York: John Wiley and Sons, Inc., 1963), II, 495-576.

In view of these difficulties, it is hardly necessary to emphasize that the formulation developed below is necessarily arbitrary, and subject to the kind of justification outlined above. However, this is not to say that the formulation could not be precisely stated or rigorously applied to the data.

Criterion variables. --- The four criterion variables tested in the hypothesis were operationally defined as follows:

1. *Size*. Of the four variables, size alone was naturally disposed to precise quantitative measurement. By "size" was meant the countable rules in a corpus. A rule in a corpus was identifiable by a given, that is, assigned notation in the cataloging code (5A(1), 5A(1a), etc.). Measurement of the variable involved a simple count of the rules in each corpus.

2. *Redundancy*. In terms of general usage, redundancy refers to unnecessary repetition or duplication. This was similarly applied in the hypothesis to mean the unnecessary repetition or duplication of conditions and specifications of entry in the rules in each corpus.

With respect to repetition, redundancy was said to occur where rules of the type:

> Enter a dramatization based on a novel ...
> under the playwright ... (ALA 23A)

and

> Enter a novelization based on a play ...
> under the novelist ... (ALA 23b)

specify not only an equivalent entry, but an equivalent set of structural descriptions. In other words, both rules specify entry under author, where author is a person categorized variably as playwright, novelist, etc., and at the same time specify equivalent requirements for the entry: novel based on drama, drama based on novel, etc. Measurement of the variable called for the assignment of structural descriptions to the rules in each corpus, and the identification of equivalency in the descriptions.

3. Complexity. By "complexity" was meant the countable requirements or conditions specified in the rules in each corpus for the determination of an entry type (author or title), and the frequency of their specification in the rules. A requirement or condition for entry was identifiable by its expression in the rules, for example the identification of "principal responsibility," "collaboration of authors," the "lack of an author," an "editor named on the title page," a "collective title," and so on. Measurement of the variable was facilitated by the structural descriptions assigned to the rules for use in determining their redundancy. On the basis of the requirements noted in the descriptions, two counts were taken: (a) the total number of requirements or conditions used by each corpus for the determination of entry, and (b) the frequency with which each requirement was specified in the corpus.

4. Specificity. By "specificity" was meant the exactness of the data, namely the requirements or conditions for entry, posited by the rules in each corpus. Conditons were exact if they permitted objective observation in contrast to subjective, intuitive interpretation. For instance, the condition of "principal responsibility" posited as such is an obscure, intuition-bound notion, subject to the interpretation and philosophical insights of a cataloger. On the other hand,

with the specification that principal responsibility is the first-named author on the title page, the name typographically or lexically prominent, etc., explication of the notion is in terms of direct observable constructs, independent of any special semantic significance or interpretability. In this sense, the condition is said to be exact. As in the case of the criterion variable of complexity, measurement of specificity was facilitated by the structural descriptions assigned to the rules for use in determining their redundancy. In particular, measurement involved determining exactness in the conditions noted in the structural descriptions, and the degree to which it characterized the conditions in each corpus of entry rules.

Since testing of the variables was contingent upon the assignment of structural descriptions to the rules, the variable of redundancy, which called for the assignment of these descriptions, was tested first in the hypothesis.

Procedures for assigning structural descriptions to the rules to test their redundancy involved the use of certain commonly accepted abstract symbols to represent the verbal statements comprising the rules. Although it was originally thought possible to test the variable using the verbal construct of the rules, examination of these constructs indicated that greater uniformity and objectivity in the measurement could be obtained if the constructs were reduced to their structural components.

The structural descriptions assigned to the corpus of AA rules are given in Table 7, and those assigned to the corpus of ALA rules are given in Table 8. Descriptions of both sets of rules are in the form: $\mathcal{X} \rightarrow \mathcal{Y}$, with the interpretation $\mathcal{X}$ "goes to" $\mathcal{Y}$. Each $\mathcal{X}$ in the description represents a bibliographic categorization specified by the rule: "work by joint authors," "anonymous work," "correspondence," etc., and is distinguished from every other $\mathcal{X}$ by certain features or structural components specified in the rules as being unique and conditional for entry of that categorization. Thus, for a "work by joint authors," two or more persons are features specified by ALA 3A as being unique to that categorization, and a condition for entry.

The features of each $\mathcal{X}$ are represented in the description by a matrix of pluses and minuses. The presence of a particular feature is symbolized by a plus, +. The absence of a feature is symbolized by a minus, -, and the presence or absence of a feature (may or may not be unique) by a combination of a plus and a minus, ±, the whole matrix of features being enclosed in brackets.[10] Thus, for a "work by joint authors," the features "two or more persons" would be represented as:

$$\begin{bmatrix} + \text{ author}_1 \\ + \text{ author}_2 \\ \pm \text{ author}_n \end{bmatrix}$$

---

[10] The use of a feature analysis of this kind is not uncommon. Perhaps, its principal use has been in the area of linguistics. See, for example: Roman Jakobson and Morris Halle, Fundamentals of Language (The Hague: Mouton & Co., 1956), pp. 3-36.

Table 7

Structural Descriptions for the Corpus of AA Rules of Entry

AA Rule	Structural Description
1A	Work of single authorship → A (where A = author = personal name) ± author + identification of authorship
2A	Work of unknown or uncertain authorship or by unnamed group → T (where T = title of work) ± { author attribution of authorship groups without a name } - identification of authorship
2C	Work of unknown or uncertain authorship or by unnamed group → T (where T = title of work) ± { author attribution of authorship groups without a name } - identification of authorship + clues to authorship { initials alphabetical devices definite article and descriptive phrase indefinite article and descriptive phrase }
3A	Work of shared authorship → Principal A (where Principal A = principal author = personal name) + { joint collaboration of authors separate contribution of authors exchange between persons shared responsibility among compilers, editors, translators, adapters, etc. } - collaboration of artist and author - collaboration of reporter and reported - collaboration of writer and nominal author - supervisory editor + author$_1$ + author$_2$ ± author$_n$ + identification of principal responsibility by means of typographical or lexical prominence of author name

Table 7 (continued)

AA Rule	Structural Description
3B1	Work of shared authorship → A (where A = author_1 = personal name) $\left[\begin{array}{l} +\left\{\begin{array}{l} \text{joint collaboration of authors} \\ \text{separate contribution of authors} \\ \text{exchange between persons} \\ \text{shared responsibility among compilers, editors,} \\ \quad \text{translators, adapters, etc.} \end{array}\right\} \\ -\ \text{collaboration of artist and author} \\ -\ \text{collaboration of reporter and person reported} \\ -\ \text{collaboration of writer and nominal author} \\ -\ \text{supervisory editor} \\ +\ \text{author}_1 \\ +\ \text{author}_2 \\ \pm\ \text{author}_3 \\ -\ \text{identification of principal responsibility by means of} \\ \quad \text{typographical or lexical prominence of author name} \end{array}\right]$
3B2	Work of shared authorship → T (where T = title of work) $\left[\begin{array}{l} +\left\{\begin{array}{l} \text{joint collaboration of authors} \\ \text{separate contributions of authors} \\ \text{exchange between persons} \\ \text{shared responsibility among compilers, editors,} \\ \quad \text{translators, adapters, etc.} \end{array}\right\} \\ -\ \text{collaboration of artist and author} \\ -\ \text{collaboration of reporter and reported} \\ -\ \text{collaboration of writer and nominal author} \\ -\ \text{supervisory editor} \\ +\ \text{author}_1 \\ +\ \text{author}_2 \\ +\ \text{author}_3 \\ +\ \text{author}_4 \\ \pm\ \text{author}_n \\ -\ \text{identification of principal responsibility by means} \\ \quad \text{of typographical or lexical prominence of} \\ \quad \text{author name} \end{array}\right]$

Table 7 (continued)

AA Rule	Structural Description
4A	Work produced under editorial direction → A (where A = editor = personal name) $\left[\begin{array}{l} +\text{ separate contributions of authors}\\ +\text{ author}_1\\ +\text{ author}_2\\ \pm\text{ author}_n\\ +\text{ supervisory editor}\\ +\text{ identification of supervisory editor on title page}\\ -\text{ publisher named in title of work}\\ +\text{ identification of primary responsibility of editor}\\ \quad\text{ for existence of work}\end{array}\right]$ Otherwise T (where T = title of work)
5A	Collection with collective title → A (where A = $\left\{\begin{array}{l}\text{editor}\\ \text{compiler}\end{array}\right\}$ = personal name) $\left[\begin{array}{l} -\text{ manuscript collections}\\ +\text{ separate pieces of author}\\ +\text{ collective title}\\ +\text{ author}_1\\ +\text{ author}_2\\ \pm\text{ author}_n\\ \pm\left\{\begin{array}{l}\text{editor}\\ \text{compiler}\end{array}\right\}\\ +\text{ identification of an }\left\{\begin{array}{l}\text{editor}\\ \text{compiler}\end{array}\right\}\end{array}\right]$ If - identification of an $\left\{\begin{array}{l}\text{editor}\\ \text{compiler}\end{array}\right\}$, Then T (where T = title of collection)
5B	Collection without collective title → A (where A = author$_1$ = personal name) Unless - author$_1$, then T (where T = title of first piece) $\left[\begin{array}{l} -\text{ manuscript groups without title}\\ -\text{ anonymous composite manuscripts}\\ +\text{ separate pieces by authors}\\ -\text{ collective title}\end{array}\right.$

Table 7 (continued)

AA Rule	Structural Description
	$\left[\begin{array}{l} + \text{ collective title page} \\ + \text{ author}_1 \\ + \text{ author}_2 \\ \pm \text{ author}_n \\ \pm \left\{\begin{array}{l}\text{editor}\\ \text{compiler}\end{array}\right\} \end{array}\right]$ If − collective title page, then A = author$_1$ in collection Unless − author$_1$, then T (where T = title of first piece in collection)
7A	Work of adapter and original author − A (where A = adapter = personal name) $\left[\begin{array}{ll} + \text{ adaption in} & \left\{\begin{array}{l}\text{different literary style:}\\ \text{paraphrase, epitome, versions for}\\ \quad\text{children}\\ \text{different literary form:}\\ \text{dramatization, novelization,}\\ \quad\text{versification}\end{array}\right\} \\ - \left\{\begin{array}{l}\text{translations}\\ \text{condensations}\\ \text{abridgments}\\ \text{enlargements}\\ \text{revisions}\end{array}\right\} & \\ \pm \text{ adapter} & \\ + \text{ identification of adapter} & \\ + \text{ original author} & \end{array}\right]$ If − identification of adapter Then T (where T = title of adaption)
8A	Work of artist and author − A (where A = $\left\{\begin{array}{l}\text{artist}\\ \text{author}\end{array}\right\}$ = first personal name) $\left[\begin{array}{l} + \text{ artist} \\ + \text{ author} \\ + \text{ collaboration of artist and author} \\ + \text{ identification of artist and author on title page} \end{array}\right]$ Unless second personal name has typographical or lexical prominence, Then A = $\left\{\begin{array}{l}\text{artist}\\ \text{author}\end{array}\right\}$ = second personal name

Table 7 (continued)

AA Rule	Structural Description
8B	Work of artist and author → A (where A = author = personal name) $\left[\begin{array}{l} +\ \text{artist} \\ +\ \text{author} \\ -\ \text{collaboration of artist and author} \end{array}\right]$
9A	Work of biographer/critic and author → A (where A = biographer/critic = personal name) $\left[\begin{array}{l} +\ \text{biographer/critic represented as author} \\ +\ \text{author} \end{array}\right]$
9B	Work of biographer/critic and author → A (where A = author = personal name) $\left[\begin{array}{l} +\ \text{biographer/critic represented as editor} \\ +\ \text{author} \end{array}\right]$
14A	Work of reviser and original author → A (where A = original author = personal name) $\left[\begin{array}{l} +\left\{\begin{array}{l} \text{condensations} \\ \text{abridgments} \\ \text{enlargements} \\ \text{revisions} \end{array}\right\} \\ +\ \text{reviser represented as editor} \\ +\ \text{original author} \end{array}\right]$
14B	Work of reviser and original author → A (where A = reviser = personal name) $\left[\begin{array}{l} +\left\{\begin{array}{l} \text{condensations} \\ \text{abridgments} \\ \text{enlargements} \\ \text{revisions} \end{array}\right\} \\ +\ \text{reviser represented as new author} \\ +\ \text{original author} \end{array}\right]$
15A	Work of translator and author → A (where A = author = personal name) $\left[\begin{array}{l} -\ \text{free translation} \\ +\ \text{translator} \\ +\ \text{author} \end{array}\right]$

Table 7 (continued)

AA Rule	Structural Description
16A	Work of writer and nominal author → A (where A = nominal author = personal name) $\begin{bmatrix} \text{+ writer} \\ \text{+ nominal author} \\ \text{+ collaboration of writer and nominal author} \end{bmatrix}$
19B	Other related works → $\left\{\begin{matrix} A \\ T \end{matrix}\right\}$ (where $\left\{\begin{matrix} A = \text{author} = \text{personal name} \\ T = \text{title of work} \end{matrix}\right\}$) [+ features of preceding rules]

Table 8

Structural Descriptions for the Corpus of ALA Rules of Entry

ALA Rule	Structural Description
2	Work by individual author → A (where A = author = personal name) $\begin{bmatrix} \pm \text{ author} \\ + \text{ identification of authorship} \end{bmatrix}$
3A	Work by joint authors → A (where A = author_1 = personal name) $\begin{bmatrix} - \text{ separate contributions of authors} \\ + \text{ author}_1 \\ + \text{ author}_2 \\ \pm \text{ author}_n \\ + \text{ identification of authors on title page} \end{bmatrix}$
3B	Work by joint authors → A (where A = author_1 = personal name) $\begin{bmatrix} - \text{ separate contributions of authors} \\ + \text{ author}_1 \\ + \text{ author}_2 \\ \pm \text{ author}_n \\ + \text{ identification of authors on title page} \\ + \text{ words on title page to the effect. "with the} \\ \quad \text{collaboration of"} \end{bmatrix}$
3E	Narrative → A (where A = narrator = personal name) $\begin{bmatrix} + \text{ narrator} \\ + \text{ reporter} \end{bmatrix}$
4A(1)	Work by collaborators → Principal A (where Principal A = $\begin{Bmatrix} \text{author}_1 \\ \text{author}_2 \\ \text{author}_n \end{Bmatrix}$ = personal name) $\begin{bmatrix} + \text{ separate contributions of authors} \\ + \text{ collaboration of authors} \\ + \text{ author}_1 \\ + \text{ author}_2 \\ \pm \text{ author}_n \\ + \text{ identification of principal responsibility} \end{bmatrix}$

Table 8 (continued)

ALA Rule	Structural Description
4A(2)	Work by collaborators → A (where A = author_i = personal name) $\begin{bmatrix} +\ \text{separate contributions of authors} \\ +\ \text{collatoration of authors} \\ +\ \text{author}_1 \\ +\ \text{author}_2 \\ \pm\ \text{author}_3 \\ -\ \text{identification of principal responsibility} \\ +\ \text{collective title} \end{bmatrix}$ If four or more authors Then T (where T = title of composite work)
4B(1)	$\left\{\begin{array}{l}\text{Debate}\\ \text{Dialogue}\\ \text{Conversation}\end{array}\right\}$ → A (where A = participant_i = personal name) $\begin{bmatrix} +\ \text{exchange between persons} \\ +\ \text{participant}_1 \\ +\ \text{participant}_2 \\ \pm\ \text{participant}_n \\ +\ \text{identification of participants on title page} \end{bmatrix}$
5A(1)	Collection with collective title → A (where A = $\left\{\begin{array}{l}\text{editor}\\ \text{compiler}\end{array}\right\}$ = personal name) $\begin{bmatrix} +\ \text{separate contributions of authors} \\ -\ \text{collaboration of authors} \\ +\ \text{author}_1 \\ +\ \text{author}_2 \\ \pm\ \text{author}_n \\ +\left\{\begin{array}{l}\text{editor}\\ \text{compiler}\end{array}\right\} \\ +\ \text{collective title} \end{bmatrix}$
5A(1a)	Collection with collective title → T (where T = title of collection) $\begin{bmatrix} +\ \text{separate contributions of authors} \\ -\ \text{collaboration of authors} \\ +\ \text{author}_1 \\ +\ \text{author}_2 \end{bmatrix}$

Table 8 (continued)

ALA Rule	Structural Description
	± author$_n$ + $\begin{Bmatrix}\text{editor}\\ \text{compiler}\end{Bmatrix}$ + collective title + $\left\{\begin{matrix}\text{minor editing}\\ \text{name of}\begin{Bmatrix}\text{editor}\\ \text{compiler}\end{Bmatrix}\text{ not in prominence}\\ \text{successive editors}\end{matrix}\right\}$
5A(1c)	Festschrift → A (where A = Publisher = $\begin{Bmatrix}\text{society}\\ \text{institution}\end{Bmatrix}$ = corporate name + separate contributions of authors + honorific - collaboration of authors + author$_1$ + author$_2$ ± author$_n$ ± editor + collective title + published by $\begin{Bmatrix}\text{society}\\ \text{institution}\end{Bmatrix}$ If not published by $\begin{Bmatrix}\text{society}\\ \text{institution}\end{Bmatrix}$ Then A = editor = personal name if name is prominent Otherwise T (where T = title of festschrift)
5A(2)	Collection without collective title → A (where A = author$_1$ = personal name) Unless - author$_1$, then T (where T = title of first piece) + separate contributions of authors - collaboration of authors + author$_1$ + author$_2$ ± author$_n$ + identification of authors on title page ± editor - collective title + collective title page If - collective title page, then A = author$_1$ in collection Unless - author$_1$, then T (where T = title of first piece in collection)

Table 8 (continued)

ALA Rule	Structural Description
5B	$\left\{\begin{array}{l}\text{Encyclopedia}\\ \text{Dictionary}\end{array}\right\} \rightarrow \text{T (where T = title of } \left\{\begin{array}{l}\text{encyclopedia}\\ \text{dictionary}\end{array}\right\}\text{)}$ $\left[\begin{array}{l} + \left\{\begin{array}{l}\text{encyclopedia}\\ \text{dictionary}\end{array}\right\}\\ \pm \text{ separate contributions of authors}\\ \pm \text{ author}_1\\ \pm \text{ author}_2\\ \pm \text{ author}_n\\ + \text{ editor}\end{array}\right]$ Unless $\left\{\begin{array}{l}\text{encyclopedia}\\ \text{dictionary}\end{array}\right\}$ is better known by name of editor Then A (where A = editor = personal name)
5C(4)	Collection of extracts → A (where A = $\left\{\begin{array}{l}\text{periodical}\\ \text{newspaper}\end{array}\right\}$ = name of serial) $\left[\begin{array}{l} + \text{ separate pieces by authors}\\ + \text{ compilation from a } \left\{\begin{array}{l}\text{periodical}\\ \text{newspaper}\end{array}\right\}\\ + \text{ author}_1\\ + \text{ author}_2\\ \pm \text{ author}_n\\ \pm \text{ editor}\end{array}\right]$ If work of single author, republished from a periodical, Then A = author = personal name Unless − identification of author, Then T (where T = title of collection)
5E(2)	Directory → A (where A = compiler = personal name) $\left[\begin{array}{l} - \text{ serial}\\ + \text{ directory}\\ + \text{ compiler}\\ + \text{ identification of compiler on title page}\end{array}\right]$ If − identification of compiler on title page, Then T (where T = title of directory)

Table 8 (continued)

ALA Rule	Structural Description
6A	Correspondence → A (where A = correspondent$_1$ = personal name) + exchange between persons + correspondent$_1$ + correspondent$_2$ ± correspondent$_n$ + identification of correspondents on title page $\pm\begin{Bmatrix}\text{editor}\\ \text{compiler}\end{Bmatrix}$
6B(1)	Correspondence → A (where A = correspondent$_1$ = personal name) + letters + author$_1$ ± author$_2$ + recipient$_1$ ± recipient$_2$ + identification of author(s) and recipient(s) on title page $\pm\begin{Bmatrix}\text{editor}\\ \text{compiler}\end{Bmatrix}$
6B(2)	Collected letters and biography of author → Principal A (where Principal A = $\begin{Bmatrix}\text{author}\\ \text{biographer}\end{Bmatrix}$ = personal name) + letters + author + biographer + identification of principal responsibility If − identification of principal responsibility, Then A = biographer = personal name
6C	Collected letters of writers → A (where A = $\begin{Bmatrix}\text{editor}\\ \text{compiler}\end{Bmatrix}$ = personal name) + letters + author$_1$ + author$_2$ ± author$_n$ ± recipient$_1$ ± recipient$_2$ $\pm\begin{Bmatrix}\text{editor}\\ \text{compiler}\end{Bmatrix}$

Table 8 (continued)

ALA Rule	Structural Description
14A	Radio script → A (where A = author = personal name $\begin{bmatrix} \pm \text{ author} \\ + \text{ identification of authorship} \end{bmatrix}$
16	Log of ship → A (where A = ship = proper name) $\begin{bmatrix} + \text{ log} \\ + \text{ name of ship} \\ \pm \text{ name of commanding officer} \end{bmatrix}$
19D	Work of art → Principal A (where Principal A = $\begin{Bmatrix} \text{artist} \\ \text{author} \end{Bmatrix}$ = personal name) $\begin{bmatrix} + \text{ artist} \\ + \text{ author} \\ + \text{ collaboration of artist and author} \\ + \text{ identification of principal responsibility} \end{bmatrix}$ If - identification of principal responsibility, Then A (where A = author = personal name)
19E	Work of art → A (where A = author = personal name) $\begin{bmatrix} + \text{ artist} \\ + \text{ author} \\ - \text{ collaboration of artist and author} \end{bmatrix}$
20	Revision → A (where A = original author = personal name) $\begin{bmatrix} + \text{ revisor} \\ + \text{ original author} \end{bmatrix}$ If reviser represents new author, Then A = reviser = personal name
21A	Translation → A (where A = author = personal name) $\begin{bmatrix} + \text{ translator} \\ + \text{ author} \end{bmatrix}$
22A	Abridgment → A (where A = original author = personal name) $\begin{bmatrix} + \begin{Bmatrix} \text{abridgments} \\ \text{epitomes} \\ \text{outlines} \end{Bmatrix} \\ + \text{ abridger} \\ + \text{ original author} \end{bmatrix}$ If - original author, Then T (where T = title of original work)

Table 5 (continued)

ALA Rule	Structural Description
22B	Adaption → A (where A = original author = personal name) $\begin{bmatrix} +\begin{Bmatrix}\text{adaptions}\\ \text{paraphrases}\\ \text{free translations}\end{Bmatrix} \\ +\ \text{adapter} \\ +\ \text{original author}\end{bmatrix}$ If − original author, Then T (where T = title of orginal work) Unless adapter represents new author, Then A (where A = adapter = personal name)
22C	Exerpts → A (where A = author = personal name) $\begin{bmatrix} +\ \text{separate pieces by an author} \\ +\ \text{compiler}\end{bmatrix}$
23A	Dramatization → A (where A = playwright = personal name) $\begin{bmatrix} -\ \text{musical drama} \\ +\ \text{playwright} \\ +\ \text{author dramatized}\end{bmatrix}$
23B	Novelization → A (where A = novelist = personal name) $\begin{bmatrix} +\ \text{novelist} \\ +\ \text{author novelized}\end{bmatrix}$
24	$\begin{Bmatrix}\text{Parody}\\ \text{Imitation}\end{Bmatrix}$ → A (where A = $\begin{Bmatrix}\text{parodist}\\ \text{imitator}\end{Bmatrix}$ = personal name) $\begin{bmatrix} +\begin{Bmatrix}\text{parodist}\\ \text{imitator}\end{Bmatrix} \\ +\ \text{author}\begin{Bmatrix}\text{parodied}\\ \text{imitated}\end{Bmatrix}\end{bmatrix}$
25A	Continuation → A (where A = author of continuation = personal name) $\begin{bmatrix} +\ \text{author of continuation} \\ +\ \text{author continued}\end{bmatrix}$

Table 8 (continued)

ALA Rule	Structural Description
25C	Supplement to periodical → A (where A = author of supplement = personal name) + monograph as supplement to periodical + author
26	Sequel → A (where A = author of sequel = personal name) + author of sequel + author sequeled
28	Concordance → A (where A = compiler of concordance = personal name) + compiler + author concordanced
32C8	Anonymous work → A (where A = author = personal name) ± author + identification of authorship ± phrase of authorship: indefinite article and descriptive phrase Unless - identification of authorship, Then T (where T = title of anonymous work)
32B	Anonymous work → T (where T = title of anonymous work) - author - identification of authorship + phrase of authorship on title page; definite article and descriptive phrase

Another feature unique to this categorization is the absence of "separate contributions", namely, the specification in the rule that the contributions of the authors cannot be separate or distinct. This feature would be added to the matrix and represented as − separate contributions of authors, and so on.

In the case of a few categorizations, the rules specify alternate features; for instance, the presence of an "editor or compiler", "adapter, paraphraser, or free translation", etc. In the structural descriptions, these alternate features are indicated in enclosed braces as, for example:

$$+ \begin{Bmatrix} \text{editor} \\ \text{compiler} \end{Bmatrix}$$

and are interpreted as: the presence of either editor or compiler.

Finally, each $\mathcal{Y}$ in the description $\mathcal{X} - \mathcal{Y}$ states a relation in the form $\mathcal{Y} = Z$, with the interpretation $\mathcal{Y}$ equals or is to be rewritten as Z as, for example, a "work by joint authors $\rightarrow \mathcal{O}$ (where $\mathcal{O}$ = author$_1$ = personal name)", that is to say, a work of joint authorship goes to an author main entry, where the author is the first one on the title page, written as a personal name. Thus, given the following author statement:

By

Mary E. Jones

and Larry L. Jones

and its agreement with the features noted, +author$_1$, +author$_2$, − separate contributions, etc., the categorization specified by the rule with these features would assign an author main entry, where the author would be the first one on the title page, "Mary E. Jones," stated as a personal name. The results of this test are given in Tables 9 and 10.

The second variable tested in the hypothesis was that of complexity, which called for the enumeration of the conditions used by each corpus for the determination of entry and the frequency with which they were specified in the corpus. To ensure maximum uniformity and objectivity in the measurement and in the comparison, the conditions featured in the structural descriptions were used for compiling the enumeration. In the case of each corpus, a condition was enumerated by its inclusion in a matrix, characterized with either a plus, minus, or plus/minus sign. The enumeration of the conditions in the corpus of AA rules and the frequency of their occurrence are given in Table 11. Those in the corpus of ALA rules are given in Table 12.

In the case of the third variable, namely, specificity, measurement involved determining exactness in the conditions on which selection of entry was based in the two sets of cataloging rules. Again, to obtain the desired uniformity and objectivity in the measurement and the overall comparison, the conditions noted in the structural descriptions were used as the data for testing the variable. The results of this test are given in Tables 13 and 14 for the corpus of AA and ALA rules respectively.

The last variable tested in the hypothesis was size, size being the countable rules in a corpus. The findings resulting from the testing of this variable and the other three criterion variables are discussed below.

The Findings

The findings which resulted from testing the four criterion variables in the hypothesis are presented below under the subheadings for the variables.

Redundancy. — The redundancy or unnecessary duplication of rules in a corpus was seen to affect the amenability of the corpus to formal characterization. Redundancy in the rules of a corpus was expected to present difficulties, specifically, in the way of deriving a common characterization that would embody all the features or conditions of the redundant rules, and at the same time permit the assignment of main entry specified by these rules. For instance, given ALA rule 23A:

> Enter a dramatization based on a novel ...
> under the playwright ...

the source of the dramatization and the dramatist are features requiring characterization. If its companion rule, 23B:

> Enter a novelization based on a play ...
> under the novelist ...

is determined redundant, as intimated earlier, it would be necessary to construct a common characterization that would embody the features and specifications of entry contained in this rule as well. A corpus with a large amount of redundancy in the rules was less likely to be characterized adequately. In comparison with a corpus having little or no redundancy, the former would be said to be considerably less amenable to simulation by the techniques proposed in the study.

To determine the amount of redundancy in each corpus, structural descriptions were assigned to the rules in the form and manner described earlier. The descriptions contained in Tables 7 and 8 for the AA and ALA rules respectively were then examined for equivalency in their features and specifications of entry. The results of this examination are reported in Tables 9 and 10 for the corpus of AA and ALA rules respectively. The contention of the hypothesis, namely, that the corpus of AA rules is less redundant than the corpus of ALA rules, had no support. As the tables show, 21 sets of rules in the AA corpus were found redundant. This was an average of 1.1 sets per rule. In contrast, 23 sets, or an average of .63 sets per rule, were found redundant in the ALA corpus. In short, the evidence showed the corpus of AA rules to be considerably more redundant than the comparable corpus of ALA rules.

A glance at Table 9 for the AA rules shows that the redundancy in rule 19B alone accounted for 18 of the 21 sets of rules that were determined structurally equivalent in the corpus. The structural description, derived from the verbal rule,

> Enter any other related work under its own author
> and/or title according to the general rules ...

indicated that the features and entry specifications were all those of the preceding rules, thus accounting for its redundancy with every other rule in the corpus. The redundancy in three other rules accounted for the remaining sets found structurally equivalent in the

Table 9

Redundancy in the Corpus of AA Rules of Entry as Shown in the Structural Descriptions

AA Rule	1A	2A	2C	3A	3B1	3B2	4A	5A	5B	7A	8A	8B	9A	9B	14A	14B	15A	16A	19B
1A																			X
2A																			X
2C																			X
3A																			X
3B1																			X
3B2																			X
4A																			X
5A																			X
5B																			X
7A																			X
8A																			X
8B																	X		X
9A																X			X
9B															X				X
14A														X					X
14B													X						X
15A												X							X
16A																			X
19B	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	

Total AA Rules of Entry	19
Sets of Redundant Rules	21
Average Number of Redundant Sets per Rule	1.1

Table 10

Redundancy in the Corpus of AIA Rules of Entry as Shown in the Structural Descriptions

AIA Rule	2	4B(1)	6A	14A	19E	20	21A	22A	22B	23A	23B	24	25A	26	28	32GR
2				X												X
4B(1)			X													
6A		X														
14A	X															X
19E							X									
20								X	X							
21A					X											
22A						X			X							
22B						X		X								
23A											X	X	X	X	X	
23B										X		X	X	X	X	
24										X	X		X	X	X	
25A										X	X	X		X	X	
26										X	X	X	X			
28										X	X	X	X	X		
32GR	X			X												

Total AIA Rules of Entry	16
Sets of Redundant Rules	21
Average Number of Redundant Sets per Rule	6.1

corpus. In contrast, Table 10 for the ALA rules shows that the redundancy in six rules accounted for the same number of sets, 18 out of 23, determined structurally equivalent in this corpus. The redundancy in five other rules accounted for the remaining sets. Although there were fewer rules that accounted for the redundancy in the corpus, the AA corpus was seen to be considerably less amenable to simulation by the techniques proposed in the study. The redundancy in the one rule in the corpus required that a common characterization be derived, which would incorporate all the features of the other rules in the corpus, and at the same time permit the assignment of the type of main entry specified by these rules.

Observation of the variable in the characterization could be expected, specifically, in the amount of characterization required for each corpus. If the characterization took the form of rules, it could be expected that the larger amount of redundancy in the AA corpus would be reflected in the number of reduced rules in the characterization. To the extent that this was observed in the characterization and was seen to affect its adequacy will be taken up in connection with the findings in Chapter V.

Complexity. --- The number of conditions and the frequency of their specification in the rules were indicative of the complexity of a corpus. A corpus with a large number of conditions could be expected to require extensive characterization in the way of description as well as computation. This complexity in a corpus was seen to have its affect on the efficiency of the characterization, specifically, in the way of accomplishing the necessary assignment of main entries.

To determine the amount of complexity in each corpus, the conditions featured in the structural descriptions were listed and the frequency of their occurrence noted. The conditions compiled from the structural descriptions for the corpus of AA rules are given in Table 11. Those from the structural descriptions for the corpus of ALA rules are given in Table 12. As can be seen, the contention of the hypothesis, namely, that the AA corpus is less complex, had no support. The tables show that there were 53, or an average of 2.7 conditions per rule in the AA corpus, which compares with 91, or an average of 2.5 conditions per rule in the ALA corpus. However, it is of interest to note that the conditions in the AA corpus were cited on an average of 2.7 times, while those in the ALA corpus were cited on an average of 1.8 times. Although it would appear that the AA corpus was more complex in this respect, a comparison of the features in the structural descriptions for comparable sets of rules shows that the corpus, on the contrary, achieved more generality by its consistent use of a single term or phrase to describe a condition. For instance, in AA rules 3A and 3b1, collaboration is featured between authors (author$_1$, author$_2$, etc.), while in such comparable rules as 3A, 6A, and 6B(1), it is featured variously between "authors", "correspondents", and "recipients".

Thus, the findings for the variable of complexity indicated that the corpus of AA and ALA rules were almost equally complex and, in this respect, almost equally amenable to simulation by the techniques proposed in the study. It could be expected that the variable would be observed in the characterization in the number of specifications

Table 11

Complexity in the Corpus of AA Rules of Entry as Shown in the Number of Conditions and Their Frequency in the Structural Descriptions

Item Number	Condition	Frequency
1.	+ adaption in {different literary style: paraphrase, epitome, version for children different literary form: dramatization, novelization, versification}	2
2.	± adapter	2
3.	- anonymous composite manuscripts	2
4.	+ artist	3
5.	+ author	6
6.	± author	2
7.	+ author$_1$	7
8.	- author$_1$	2
9.	+ author$_2$	7
10.	+ author$_3$	2
11.	± author$_3$	2
12.	+ author$_4$	2
13.	± author$_n$	6
14.	± {author attribution of authorship groups without a name}	3
15.	+ biographer represented as author	2
16.	+ biographer represented as editor	2

Table II (continued)

Item Number	Condition	Frequency
17.	+ clues to authorship {initials alphabetical devices definite article and descriptive phrase indefinite article and descriptive phrase}	2
18.	+ collaboration of artist and author	2
19.	- collaboration of artist and author	5
20.	- collaboration of reporter and reported	4
21.	+ collaboration of writer and nominal author	2
22.	- collaboration of writer and nominal author	4
23.	+ collective title	2
24.	- collective title	2
25.	+ collective title page	2
26.	- collective title page	2
27.	+ {condensations abridgments revisions}	2
28.	± {editor compiler}	3
29.	- free translations	2
30.	+ identification of adapter	2
31.	- identification of adapter	2
32.	+ identification of an {editor compiler}	2

Table II (continued)

Item Number	Condition	Frequency
33.	- identification of an {editor compiler}	2
34.	+ identification of artist and author on title page	2
35.	+ identification of authorship	3
36.	- identification of authorship	3
37.	+ identification of primary responsibility of editor for existence of work	2
38.	+ identification of principal responsibility by means of typographical or lexical prominence of author name	3
39.	- identification of principal responsibility by means of typographical or lexical prominence of author name	3
40.	+ identification of supervisory editor on title page	2
41.	+ {joint collaboration of authors separate contribution of authors exchange between persona shared responsibility among compilers, editors, translators, adapters, etc.}	4
42.	- manuscript collections	3
43.	+ nominal author	2
44.	+ original author	4
45.	- publisher named in title page	2
46.	+ reviser represented as editor	2
47.	+ reviser represented as new author	2

Table II (continued)

Item Number	Condition	Frequency
48.	+ separate pieces by authors	3
49.	+ supervisory editor	2
50.	- supervisory editor	4
51.	- {translations condensations abridgments enlargements revisions}	2
52.	+ translator	2
53.	+ writer	2

Total AA Rules of Entry 19
Average Number of Conditions
per Rule 2.7
Average Frequency of Conditions 2.7

Table 12

Complexity in the Corpus of ALA Rules of Entry as Shown in the Number of Conditions and Their Frequency in the Structural Descriptions

Item Number	Condition	Frequency
1.	+ abridger	1
2.	$+\begin{Bmatrix}\text{abridgments}\\ \text{epitomes}\\ \text{outlines}\end{Bmatrix}$	1
3.	$+\begin{Bmatrix}\text{adaptions}\\ \text{paraphrases}\\ \text{free translations}\end{Bmatrix}$	1
4.	+ adapter	1
5.	+ adapter as new author	1
6.	+ artist	2
7.	+ association of editor's name with encyclopedia or dictionary	1
8.	+ author	5
9.	- author	1
10.	± author	3
11.	+ author$_1$	12
12.	± author$_1$	1
13.	+ author$_2$	11
14.	± author$_2$	2
15.	+ author$_3$	1
16.	± author$_3$	1

Table 12 (continued)

Item Number	Condition	Frequency
17.	+ author_4	1
18.	± author_n	10
19.	+ author concordanced	1
20.	+ author continued	1
21.	+ author dramatized	1
22.	+ author novelized	1
23.	+ author of continuation	1
24.	+ author of sequel	1
25.	+ author $\left\{\begin{matrix}\text{parodied}\\ \text{imitated}\end{matrix}\right\}$	1
26.	+ author sequeled	1
27.	+ biographer	1
28.	+ collaboration of artist and author	1
29.	- collaboration of artist and author	1
30.	+ collaboration of authors	2
31.	- collaboration of authors	4
32.	+ collective title	4
33.	- collective title	1
34.	+ collective title page	1
35.	- collective title page	1
36.	+ compilation from $\left\{\begin{matrix}\text{periodical}\\ \text{newspaper}\end{matrix}\right\}$	2
37.	+ compiler	3

Table 12 (continued)

Item Number	Condition	Frequency
38.	+ correspondent$_1$	1
39.	+ correspondent$_2$	1
40.	± correspondent$_n$	1
41.	+ directory	1
42.	+ editor	1
43.	± editor	3
44.	$+\begin{Bmatrix}\text{editor}\\ \text{compiler}\end{Bmatrix}$	3
45.	$\pm\begin{Bmatrix}\text{editor}\\ \text{compiler}\end{Bmatrix}$	2
46.	$+\begin{Bmatrix}\text{encyclopedia}\\ \text{dictionary}\end{Bmatrix}$	1
47.	+ exchange between persons	2
48.	+ honorific	1
49.	+ identification of authors and recipients on title page	1
50.	+ identification of authors on title page	3
51.	+ identification of authorship	3
52.	- identification of authorship	3
53.	+ identification of compiler on title page	1
54.	- identification of compiler on title page	1
55.	+ identification of correspondents on title page	1
56.	+ identification of participants on title page	1

Table 12 (continued)

Item Number	Condition	Frequency
57.	+ identification of principal responsibility	3
58.	- identification of principal responsibility	3
59.	+ letters	3
60.	+ log	1
61.	$+\left\{\begin{array}{l}\text{minor editing}\\ \text{name of } \left\{\begin{array}{l}\text{editor}\\ \text{compiler}\end{array}\right\} \text{ not in prominence}\\ \text{successive editors}\end{array}\right\}$	1
62.	+ monograph as supplement to periodical	1
63.	- musical drama	1
64.	± name of commanding officer	1
65.	+ name of ship	1
66.	+ narrator	1
67.	+ novelist	1
68.	+ original author	3
69.	- original author	2
70.	$+\left\{\begin{array}{l}\text{parodist}\\ \text{imitator}\end{array}\right\}$	1
71.	+ participant$_1$	1
72.	+ participant$_2$	1
73.	± participant$_n$	1
74.	+ phrase of authorship on title page definite article and descriptive phrase	1

Table 12 (continued)

Item Number	Condition	Frequency
75	± phrase of authorship: indefinite article and descriptive phrase	1
76	+ playwright	1
77	+ published by $\left\{\begin{array}{l}\text{society}\\ \text{institution}\end{array}\right\}$	1
78	- published by $\left\{\begin{array}{l}\text{society}\\ \text{institution}\end{array}\right\}$	1
79	+ recipient_1	1
80	± recipient_1	1
81	± recipient_2	2
82	+ reporter	1
83	+ reviser	1
84	+ reviser as new author	1
85	+ separate contributions, pieces of an author	2
86	+ separate contributions, pieces of authors	8
87	- separate contributions of authors	1
88	± separate contributions of authors	1
89	- serial	1
90	+ translator	1
91	+ words to the effect: "with the collaboration of"	1
Total ALA Rules of Entry	36	
Average Number of Conditions per Rule	2.5	
Average Frequency of Conditions	1.8	

or nodes in the trees and the number of steps in the trees that are needed to satisfy the entry assignment specified by the verbal rules. With a comparable average number of conditions in the two sets of verbal rules, it could be expected then that there would be a comparable average number of nodes and steps in the trees developed for the rules. The findings with respect to both characterizations are presented in Chapter V.

Specificity. — The number of conditions in the rules that were exact (permitting objective observation in contrast to subjective, intuitive interpretation) was considered indicative of the specificity of a corpus. Exact conditions were seen to accommodate formal characterization by virtue of their physical display on the title page (a translator or an artist observed by those words or the terms "translated by" and "illustrated by", principal responsibility observed as the first author name, the name typographically prominent, etc.). In this respect, a corpus with a large number of such conditions is said to be more amenable to simulation by the proposed techniques than one with a small number of such conditions.

To determine the specificity of each corpus, the conditions noted in the structural descriptions for the rules were examined for their physical display in the title page data. A condition permitting objective observation was categorized as exact. In contrast, a condition permitting subjective, intuitive interpretation was categorized as inexact. The results of this measurement are given in Tables 13 and 14, for the AA and ALA corpus respectively. As can be seen, there was

Table 13

Specificity in the Conditions of the Structural Descriptions for the Corpus of AA Rules of Entry

Item Number	Condition	Exact	Inexact
1.	+ adaption in: different literary style: paraphrase, epitome, version for children	x	
	different literary form: dramatization, novelization, versification	x	
2.	+ adapter	x	
3.	- anonymous composite manuscripts		x
4.	+ artist	x	
5.	+ author	x	
6.	± author	x	
7.	+ author_1	x	
8.	- author_1	x	
9.	+ author_2	x	
10.	+ author_3	x	
11.	± author_3	x	
12.	+ author_4	x	
13.	± author_n	x	
14.	± author	x	
	attribution of authorship	x	
	groups without a name	x	
15.	+ biographer represented as author	x	
16.	+ biographer represented as editor	x	

Table 13 (continued)

Item Number	Condition	Exact	Inexact
17.	+ clues to authorship initials	x	
	alphabetical devices	x	
	definite article and descriptive phrase	x	
	indefinite article and descriptive phrase	x	
18.	+ collaboration of artist and author	x	
19.	- collaboration of artist and author	x	
20.	- collaboration of reporter and reported	x	
21.	+ collaboration of writer and nominal author	x	
22.	- collaboration of writer and nominal author	x	
23.	+ collective title		x
24.	- collective title		x
25.	+ collective title page		x
26.	- collective title page		x
27.	+ condensations	x	
	abridgments	x	
	revisions	x	
28.	± editor	x	
	compiler	x	
29.	- free translations	x	
30.	+ identification of adapter		x
31.	- identification of adapter		x
32.	+ identification of an editor		x
	compiler		x
33.	- identification of an editor		x
	compiler		x

Table 15 (continued)

Item Number	Condition	Exact	Inexact
34.	+ identification of artist and author on title page	x	
35.	+ identification of authorship		x
36.	- identification of authorship		x
37.	+ identification of primary responsibility of editor for existence of work		x
38.	+ identification of principal responsibility by means of typographical or lexical prominence of author name	x	
39.	- identification of principal responsibility by means of typographical or lexical prominence of author name	x	
40.	+ identification of supervisory editor on title page	x	
41.	+ joint collaboration of authors	x	
	separate contributions of authors		x
	exchange between persons	x	
	shared responsibility among compilers, editors, translators, adapters, etc.	x	
42.	- manuscript collections		x
43.	+ nominal author	x	
44.	+ original author		x
45.	- publisher named in title of work	x	
46.	+ reviser represented as editor	x	
47.	+ reviser represented as new author	x	
48.	+ separate contributions, pieces of author		x
49.	+ supervisory editor	x	

Table 13 (continued)

Item Number	Condition	Exact	Inexact
50.	- supervisory editor	x	
51.	- translations	x	
	condensations	x	
	abridgments	x	
	enlargements	x	
	revisions	x	
52.	+ translator	x	
53.	+ writer	x	
	Total	37	16

Total AA Rules of Entry	19
Average Number of Exact Conditions per Rule	1.9
Average Number of Inexact Conditions per Rule	.84

Table 14

Specificity in the Conditions of the Structural Descriptions for the Corpus of ALA Rules of Entry

Item Number	Condition	Exact	Inexact
1.	+ abridger	x	
2.	+ abridgements epitomes outlines	x x x	
3.	+ adaptions paraphrases free translations	x x x	
4.	+ adapter	x	
5.	+ adapter as new author		x
6.	+ artist	x	
7.	+ association of editor's name with encyclopedia or dictionary		x
8.	+ author	x	
9.	- author	x	
10.	± author	x	
11.	+ author$_1$	x	
12.	± author$_1$	x	
13.	+ author$_2$	x	
14.	± author$_2$	x	
15.	+ author$_3$	x	
16.	± author$_3$	x	
17.	+ author$_4$	x	

Table 14 (continued)

Item Number	Condition	Exact	Inexact
18.	± author$_n$	x	
19.	+ author concordanced	x	
20.	+ author continued	x	
21.	+ author dramatized	x	
22.	+ author novelized	x	
23.	+ author of continuation	x	
24.	+ author of sequel	x	
25.	+ author {parodied imitated}	x x	
26.	+ author sequeled	x	
27.	+ biographer	x	
28.	+ collaboration of artist and author	x	
29.	- collaboration of artist and author	x	
30.	+ collaboration of authors		x
31.	- collaboration of authors		x
32.	+ collective title		x
33.	- collective title		x
34.	+ collective title page		x
35.	- collective title page		x
36.	+ compilation from {periodical newspaper}	x x	
37.	+ compiler	x	
38.	+ correspondent$_1$	x	

Table 14 (continued)

Item Number	Condition	Exact	Inexact
39.	+ correspondent$_2$	x	
40.	± correspondent$_n$	x	
41.	+ directory	x	
42.	+ editor	x	
43.	± editor	x	
44.	+ {editor compiler}	x x	
45.	± {editor compiler}	x x	
46.	+ {encyclopedia dictionary}	x x	
47.	+ exchange between persons	x	
48.	+ honorific	x	
49.	+ identification of author and recipients on title page	x	
50.	+ identification of authors on title page	x	
51.	+ identification of authorship		x
52.	- identification of authorship		x
53.	+ identification of compiler on title page	x	
54.	- identification of compiler on title page	x	
55.	+ identification of correspondents on title page	x	
56.	+ identifcation of participants on title page	x	
57.	+ identification of principal responsibility		x

Table 14 (continued)

Item Number	Condition	Exact	Inexact
58.	- identification of principal responsibility		x
59.	+ letters	x	
60.	+ log	x	
61.	+ minor editing	x	
	name of editor compiler not in prominence		x
	successive editors		x
62.	+ monograph as supplement to periodical	x	
63.	- musical drama		x
64.	± name of commanding officer	x	
65.	+ name of ship	x	
66.	+ narrator	x	
67.	+ novelist	x	
68.	+ original author		x
69.	- original author		x
70.	+ parodist	x	
	imitator	x	
71.	+ participant$_1$	x	
72.	+ participant$_2$	x	
73.	± participant$_n$	x	
74.	+ phrase of authorship on title page: definite article and descriptive phrase	x	
75.	± phrase of authorship. indefinite article and descriptive phrase	x	

Table 14 (continued)

Item Number	Condition	Exact	Inexact
76.	+ playwright	x	
77.	+ published by {society institution}	x x	
78.	- published by {society institution}	x x	
79.	+ recipient$_1$	x	
80.	± recipient$_1$	x	
81.	± recipient$_2$	x	
82.	+ reporter	x	
83.	+ reviser	x	
84.	+ reviser as new author	x	
85.	+ separate contributions, pieces of an author		x
86.	+ separate contributions, pieces of an author		x
87.	- separate contributions of authors		x
88.	± separate contributions of authors		x
89.	- serial		x
90.	+ translator	x	
91.	+ words to the effect "with the collaboration of"	x	
	Total	70	21

Total ALA Rules of Entry	36
Average Number of Exact Conditions Per Rule	1.9
Average Number of Inexact Conditions Per Rule	.58

no difference in the two sets of rules for this variable. The contention of the hypothesis that conditions in the AA corpus were enumerated more specifically had no support. Of the 53 conditions in the structural descriptions for the AA corpus, 37 or an average of 1.9 conditions per rule were determined to be exact. This compared with 70 of the 91 conditions, or an average of 1.9 conditions per rule, in the ALA corpus.

In the characterization, it could be expected that the variable would be observed in the specifications contained at the nodes and roots of the trees developed for each corpus. Exact conditions in the rules would be characterized as the specifications in the trees. Inexact conditions, on the other hand, could be expected to lack characterization. With a comparable average number of exact conditions in the rules, it could be expected then that there would be a comparable average number of specifications in the two characterizations. The findings with respect to both characterizations are presented in Chapter V.

Size. --- Finally, the corpus with a small number of rules was considered to be more amenable to formal characterization than one with a large number of rules. A corpus with a small number of rules was seen to require considerably less characterization. From the viewpoint of practical application and efficiency, the economy achieved in the characterization of the corpus was seen to facilitate speed and accuracy in application. As previously noted, 19 rules of entry were selected from the AA code for characterization in the study. To provide characterization for comparable rules in the ALA code, 36 rules of entry were selected. In short, there were almost twice as many rules of entry in the latter code to be characterized. The findings clearly supported the contention of the hypothesis that the corpus selected from the AA code would contain fewer rules than a counterpart corpus selected from the ALA code.

It was expected that the differences noted in the two sets of main entry rules would be reflected in the two characterizations developed in the study. If, for instance, the characterization were to take the form of rules using the descriptive devices noted earlier, these differences would be reflected in the number of rules developed in the characterization of each corpus. Observation of the variable in the characterizations and its effect, if any, on their accuracy and speed in application are discussed in connection with the findings in Chapter V and VI.

Summary

The procedures used for selecting and testing the amenability of the AA and ALA rules of entry to formal characterization were discussed in detail in this chapter. A corpus of 19 AA rules and a counterpart corpus of 36 ALA rules were selected for formal characterization. The amenability of each corpus of entry rules to formal characterization was tested in terms of the criterion variables of size, complexity, redundancy, and specificity stated in the hypothesis. With the exception of size, the findings for the variables measured in the two sets of main entry rules did not support the contentions of the hypothesis. On the contrary, the findings for the variables of complexity and specificity indicated that both sets of rules were equally or almost equally amenable to formal characterization, while the findings for the variable of redundancy clearly supported the amenability of the ALA corpus.

However, it was noted that the relevancy of these variables as factors of amenability would have to be made a posteriori, that is, in view of the characterizations developed for the two sets of rules and the findings resulting from the test of accuracy applied to these characterizations. For an examination of their relevancy, see Chapter V (pp. 237-44).

CHAPTER IV

FORMULATION OF THE SIMULATION MODELS

In the preceding chapter, the research method used to test the hypothesis was discussed in detail. A corpus of rules of entry was selected from the AA and ALA cataloging codes for formal characterization in the study, and the degree of amenability of each corpus to simulation by automatic techniques was measured in terms of the criterion variables stated in the hypothesis. The present chapter will deal with the formal characterization of the rules of entry selected from the two cataloging codes. The construction of the formal characterization will be discussed, and the simulation models comprising the characterization of each corpus will be presented.

The first section will describe the methods and procedures used in formulating the characterization, and in designing the simulation models for each corpus. The second section will present the simulation models developed in the study. The adequacy of the simulation models to generate main entry choices accurately and automatically will be taken up in Chapters V and VI.

Formulation of the Characterization and Design of the Simulation Models

The characterization of the rules of entry selected from the two cataloging codes is a central concern of the pages that follow. Before discussing the methods and procedures used in the formulation of this characterization, it is well to recapitulate at this point the fundamental aims and objectives of the characterization, and to briefly enumerate the limitations imposed on its formulation.

The fundamental aim of the characterization of each corpus is a formalized model of entry selection that will simulate the human ordering of this environment and at the same time meet certain conditions of adequacy and simplicity to provide justification for the choice of one formalized model of entry selection over another. In essence, the task of characterization is to make explicit a process that every cataloger performs implicitly. It is expected that the characterization will, for instance, represent the structure and content (the principles and rules) governing the entry selection process, and that it will incorporate in this representation certain capacities that the cataloger brings to bear in performing the process. First, it must be capable of recognizing certain commonly acknowledged items of information comprising the bibliographic data on which selection of entry is based, and second, on the basis of this recognition, it must be capable of determining what is an appropriate main entry choice. These capacities are, of course, essential to the characterization if it is to accomplish its main objective, namely to demonstrate that selection of entry can be accomplished in a completely objective, mechanical manner without the use of human intuition and foreknowledge.

With respect to the formulation of the characterization, certain constraints were imposed. First, the formal characterization of the entry process developed in the study was not intended to constitute a model of the intellectual or intuitive processes of the cataloger performing the process. As was pointed out in Chapter II, the fact that catalogers, for instance, can comprehend an unlimited variety of title page data, and can judge and determine their usefulness in formulating an entry choice immediately indicates that their capacities are quite different from an automaton that performs operations on the basis of printed characters and symbols presented to it. To say that the characterization then simulates the entry process is not to say that the catalogers who perform the process accomplish it using the same strategies, etc. of the characterization. To characterize the process is one task. To characterize the cataloger's performance of it is another. The two problems are obviously related but not identical.

Second, the economic or technical feasibility of the characterization was eliminated from consideration in the study. While the importance of these issues was not precluded, the exploratory nature of the investigation did not warrant their consideration. Consequently, the methods developed in the study did not provide a computer program, but rather the conceptual base and provisionary steps necessary before programming and technical feasibility can be considered.

Finally, with respect to its operational capacity, the characterization was limited to the identification and manipulation of cataloging data on the title pages of English language books only. Cataloging data on the title page consisted specifically of title and author statements. The use of title page data as the primary and major source for determining entry is openly acknowledged in the AA code, which states that

> the entry for a work is normally based on the statements that appear on the title page, or on any part of the work that is used as its substitute, but other statements that are openly expressed . . . are also taken into account. Material that appears only in the preface, introduction, or text enters into the determination of the main entry only when the information on the title page, etc., is ambiguous or insufficient.[1]

In the ALA code, sources for determining entry are not given. However, the use of title page data is tacitly assumed, if its priority is not. (In the case of six of the ALA rules selected for characterization, it is designated as the sole source of information on which selection of entry is based; namely, ALA rules 3A, 3B, 4B(1), 5A(2), 5E(2), and 6A). Although it was recognized that the restriction to the use of title page data was severe, it was prompted out of necessity. In no way was it meant to infer that selection of entry can or should be accomplished solely on the use of title page data.

The characterization of each corpus of rules was developed in terms of the components that serve to identify the entry process and the devices that are available to the performance of the process. In the case of both codes, provision of entry was identified as comprising two fundamental components, 1) the recognition of the presence of authorship in a work, and on the basis of this recognition, 2) the identification of author responsibility for a work. The devices available to the performance of the entry process were identified as:

---

[1] The American Library Association and others, Anglo-American Cataloging Rules, North American Text (Chicago: American Library Association, 1967), p. 9.

1) the rules of entry contained in each code, and 2) the bibliographic data used by those rules to determine the main entry choice. The method of characterizing each of these components using the devices available to them is discussed below under two subheadings. In the first subheading, the recognition component is discussed. This is followed by a second subheading, in which the selection component is discussed.

1. The recognition component. --- Recognizing the presence of an author on the title page of a book is a task that every cataloger performs intuitively. Indeed, it is so fundamental that instruction in the task is not necessary. Mere visual observation of the bibliographic data is sufficient to inform the cataloger as to the presence or absence of this item. Yet, it is precisely the cataloger's visual observation of the data that the characterization must represent.

To give an account of the cataloger's ability to recognize the presence of an author on the title page is actually to describe his capacity to recognize a "name". Although an author is described in both codes as being a person or a corporate body, the cataloger actually identifies the person or corporate body as a name. There is perhaps no way of stating precisely how the cataloger establishes the identity of a "name" on the title page, but it seems evident that in addition to perception, intuition, and certain other native capacities, he is greatly assisted by the physical and syntactic display of the bibliographic data. By physical is meant the position, size, and composition (alphabet, numbers, etc.) of the data and their distribution on the page. By syntactic is meant those lexical units, words, of which the data are composed and the manner of their combination on the title page. The way in which the physical and syntactic display of the bibliographic data can be thought to assist the cataloger is in its patterning, which furnishes the cataloger with certain cues for recognizing the presence of an author name when there is one on the title page. Examples of this patterning are the three title page displays given below.

(1)

THE YEAR
OF THE WHALE
By Victor B. Scheffer

(2)

PRINCIPLES OF
MATHEMATICAL LOGIC
by
D. HILBERT AND W. ACKERMAN

(3)

THIRTY YEARS OF CHAMPIONSHIP GOLF
BY DAI REES
with John Ballantine

Examination of these displays, typical of the title pages in many English language books, reveals striking similarities in the physical arrangement and distribution of the data, and in the use of certain syntactic units embedded in the data. In terms of arrangement and distribution, the data in the three displays are seen to comprise two physically distinct statements or strings of syntactic units, namely, a title statement exhibited as the initial or first string in the display, and an author statement exhibited as the sub or second string in the display. This similarity in the position and sequence of the data was evidenced by an extensive examination of the title pages of hundreds of books, and strongly suggested that a cue the cataloger (or a characterization) might use in the recognition of authorship was the physical patterning of the data themselves. For instance, given the patterning above, the characterization would recognize the author name as being the sub or second string of units in the displays. In a patterning where this string was absent, for example:

(4) THE ART AND SCIENCE
OF HUMAN MOVEMENT

there would, of course, be no recognition. Certain complexities are introduced in author name recognition when the patterning displays a reversal in the sequence of data; that is, where the author statement is the initial string and the title statement, the substring; for example:

(5) MARTIN JOOS
THE FIVE CLOCKS

and when the patterning exhibits a conjointed title-author statement, where the author statement is embedded in the title statement, and the whole is identified as an initial string. For example:

(6) THE LETTERS
OF
PFC RICHARD E. MARKS, USMC

These are complexities that must certainly be considered in constructing the characterization. From these and other cases, it seems clear that the physical patterning of the title page data alone will be insufficient as a means of recognizing all the author names that a title page may display. However, when used in conjunction with the syntactic display of the data, it provides an important cue for the characterization, and perhaps for the cataloger as well.

In terms of the syntactic display of the data, the previous title pages, with the exception of (4) and (5), exhibit the use of certain syntactic units to signal the presence of authorship. For instance, in the case of (1), (2), and (3), the syntactic unit BY is seen to signal the presence of an author name. In (2) and (3), the units AND and WITH are used to signal the presence of additional author names in these displays. In (6), the unit OF is found in conjunction with the author name. Similarly, in the case of BETWEEN, FROM, and "S.S.", the presence of an author name or names is indicated, as for example.

(7) The Correspondence
Between Benjamin Harrison and James G. Blaine

(8) LETTERS FROM WILLIAM CORBETT
TO EDWARD THORNTON

(9) CARDINAL NEWMAN'S
BEST PLAIN SERMONS

In an examination of hundreds of title pages, a regularity in the use of these seven syntactic units to signal the presence of authorship was clearly discernible. The common association of these syntactic units with author names suggested another cue that the cataloger, and a characterization, might use in the recognition of authorship. As with the physical patterning of the data, certain complexities occur that must surely be taken into consideration in constructing the characterization. An obvious complexity is the variety of contexts in which these units are used. Take, for example, the unit BY which can function to express a relation in space, as in "*by* the door", a relation in time, as in "*by* the hour", a direction of movement, as in "we went *by* Champaign", a manner or mode as in "it grows dark *by* degrees", and so on. From these few examples, it is apparent that certain restrictions must be imposed on its recognition as an author name marker. This is similarly true of the other units.

Another complexity is introduced when the syntactic units are omitted in the author statement. These omissions occur largely with the units BY and AND, where they are most certainly implied but physically absent in the author statement. For example:

(10) THE HIDDEN DIMENSION
Edward T. Hall

(11) FORMAL LANGUAGES AND THEIR RELATION TO AUTOMATA

JOHN E. HOPCROFT
Cornell University, Ithaca, New York

JEFFREY D. ULLMAN
Bell Telephone Laboratories, Murray Hill, N.J.

In such cases, the physical patterning of the data can provide a particularly useful cue in making the necessary identification, namely, by identifying authorship as the substrings in the displays. While these complexities introduce a number of difficulties in constructing a characterization, the association of these syntactic units with author names is seen to provide a powerful cue in author name recognition.

The characterization of the recognition component developed in the study was formulated in terms of the concept described above. The construction of the characterization based on this concept can now be considered in greater detail.

For each corpus of main entry rules, the recognition component was characterized as a set of rules constructed in such a way that when applied directly to the title and author statements on the title page, the author name or names are identified as to their presence or absence on the title page.

In the case of each corpus, the rules were constructed out of a set of seven recognition markers, namely, those syntactic units identified earlier in the discussion as signalling devices for the presence of authorship BY, AND, WITH, OF, BETWEEN, FROM, 'S,S'. At the first level of development, these rules were simple instructions of the form. Recognize (that is, identify) such and such a marker when it occurs on the title page. For example, recognize BY when it occurs on the title page, and so on for each marker. In applying the rule for BY, for instance, to the title page below,

(12)

SOCRATES

by
A.E. Taylor

the following recognition would take place:

by A.E. Taylor

As noted earlier in the discussion, a difficulty is immediately encountered when general rules of this type are applied to a title page, such as the following:

(13) AIR POLLUTION BY METALLURGICAL INDUSTRIES

By
Alan Bell and John L. Sullivan

where the recognition results in:

BY METALLURGICAL INDUSTRIES

By Alan Bell and John L. Sullivan

In view of the variety of contexts in which BY and the other recognition markers can function, it was necessary that the rules place certain restrictions on their recognition. One restriction involved specifying the position in which the recognition marker must occur on the title page. Using BY as an example, the rules were constructed in the form of: Recognize BY only when it occurs in the author statement, the author statement being identified by the rules as either the initial (first) or sub (second) string of lexical units or words on the title page. In the characterization developed in the study, the author statement was identified by the rules as being the substring, thus making the rule read: Recognize BY only when it occurs in the substring. In applying this form of rule to example (13) above, the statements would be identified as:

Initial string: AIR POLLUTION BY METALLURGICAL INDUSTRIES

Substring: By
Alan Bell and John L. Sullivan

and the resulting recognition would be that of the author name: By Alan Bell. To recognize the joint author, a similar rule for the marker AND would be constructed, taking into account its particular position in the substring. For example: Recognize AND only when it occurs after BY in the substring which, when applied to the above, would produce the needed recognition: and John L. Sullivan. With rules of this kind, non-author recognitions such as the above could be successfully blocked.

However, a complication that does develop in connection with the application of these rules occurs when the author statement is embedded in the title statement and the whole is, of course, identified by the rules as an initial string. This complication arises with all but one of the recognition markers, namely, OF, FROM, BETWEEN, WITH, AND, and 'S,S'. For example:

(14) The Collected Works of C. Graham Botha

(15) LETTERS FROM SAINT AUGUSTINE
TO HIS MOTHER

(16) Conversations
Between BERNARD BERENSON and OSCAR MAHER

Additional examples are (7), (8), and (9) on pages 110 and 111.

With the application of these rules, "non" or no recognition results since there is no identification of a substring. To facilitate recognition in these cases, rules were first constructed, which permitted recognition of these markers in the initial string in the absence of a substring. While these rules produced the needed recognitions, they were also found to produce non-author recognitions when applied to title pages of the type below.

(17) THE ART AND SCIENCE OF HUMAN MOVEMENT

(18) Best Recipes From The Deep South

(19) THE WAR BETWEEN THE STATES AN OUTLINE

This difficulty necessitated that the rules place a further restriction on the recognition of these markers namely, by specifying the environment in which they must occur in the initial string For instance, in the case of four of these markers OF, FROM, BETWEEN, and WITH, it was found that their association with an author name in the initial string occurred in the context of certain lexical units, such as "conversations," "diary," "essays," "letters," "poems," "works," etc., a few examples of which are given in (14), (15), and (16) above. From such patterning, rules were constructed in the form Recognize OF etc., when it occurs in the environment of conversations, diary, essays, etc. With the construction of these rules, recognitions in cases such as (17), (18), and (19) were blocked. At the same time, the rules permitted the operation of other rules that identified these markers in the substring when they occurred. For example, given the title page.

(20)

THE COLLECTED ESSAYS
OF
JOHN F. BECK

Selected and compiled
by the Editors of Organic Farming

these rules would recognize: OF JOHN F. BECK, and allow another rule for the marker: Recognize OF, etc., in the substring after BY, etc., to make the identification of the marker in that position: of Organic Farming.

In the case of the recognition markers AND and 'S.S', the rules were unable to specify a particular environment for their recognition in the initial string. While the recognition of AND was found to be unnecessary, a point discussed later in this section, special rules had to be constructed for the recognition of 'S.S'. A simple but special rule for this marker was first in the form: Recognize _______'S.S' when it occurs in the initial string in the absence of a substring. With this rule, recognitions in (9) and similar displays could be successfully produced. However, when applied to the following:

(21) A DIETER'S CALORIE COUNTER

(22) BRENDAN BEHAN'S
NEW YORK

Drawings by
Paul Hogarth

(23) WILLE AND MARTIN'S
MERCANTILE LAW OF SOUTH AFRICA

By
Philip Millin and George Wille

either a non-author recognition, as in (21), or a non-recognition, as in (22) and (23) resulted.

For patternings such as (21), which appear to be few for this marker, non-author recognitions could not be avoided. However, one way in which the rule could assist in blocking a faulty selection later was to have it recognize the association of the marker with the pattern A ____________'S,S'. Since the A (indefinite article) was always found to introduce the marker's association with a non-author, the rule, by specifying its inclusion in the recognition, could assist in subsequently blocking the inappropriate selection of an author main entry. To facilitate recognition of this pattern, the rule was modified to: Recognize (A) __________'S,S' when it occurs in the initial string, etc., where the A is an indefinite article that may or may not occur in the pattern, indicated by the parenthesis, ( ). By enclosing the article in parentheses, the rule functions to recognize either A __________'S,S' or ____________'S,S'.

For such patternings as (22) and (23) above, where the marker is associated with an author name in the initial string and followed by a substring, it was necessary to construct two special rules that would take into account the association of the marker with an author name(s) in the initial string and its occurrence in the environment of a substring. These rules were constructed in the form: Recognize (A) ____________'S,S' when it is in the environment of drawings, forword, etc. BY, and Recognize (A) ____________$\overset{\&}{\text{AND}}$ _____________'S,S', when it is in the environment of BY, both of which would produce appropriate recognitions for (22), (23), and similar displays. The inclusion of the indefinite article in the rules was similarly for the purpose of subsequently blocking inappropriate entry selections as, for example, in (24) below.

(24) A COOK'S KITCHEN

Written and Illustrated By
Lois Royce

For two of the markers, namely WITH and AND, the rules had to account for certain exceptions to their recognition in the substring. In the case of both markers, their association with an author name occurred typically after BY:

(25) BUCHER: MY STORY

By

Lloyd M. Bucher

With Mark Rascovich

(26) QUANTITATIVE ANALYSIS

By Earl W. Kendall
In Collaboration With Morris Lawrence

(27) THE LESSONS OF HISTORY

By

Will and Ariel Durant

By constructing rules, such as: Recognize WITH, Recognize AND after BY in the substring, author name recognitions in these cases could be successfully produced. In more complex patternings, the rules were unsatisfactory. Consider, for example, the application of this rule for WITH:

(28) THE PRINCE AND THE PAUPER
A Tale for Young People of All Ages
BY
MARK TWAIN
With Illustrations by Robert Hodgson

(29) THE ARCHITECTURE OF THE BONE
BY
Michael C. Hall
WITH THE ASSISTANCE OF DOROTHY S. KINOSHITA

where the recognitions for the marker, indicated by underscoring, are: (28) With Illustrations, and (29) WITH THE ASSISTANCE. The occurrence of the marker in a medial position between two or more BYs as in (28), and in conjunction with the marker OF in (29) was typical of many title page displays. In almost all such cases, the marker appeared in association with a non-author. This necessitated that the rules provide for some exceptions to the general method of author name recognition. These exceptions were built into the rules as, for instance, in the case of WITH. Recognize WITH after BY in the substring except when (a) WITH is in the environment of BY ... WITH ... BY, (b) WITH is in the environment of BY ... WITH ... OF. Similarly, in the case of AND, exceptions were built into the rule for this marker in the form of certain environmental restrictions. For instance, to avoid non-author recognition in these more complex patternings:

(30) MOTHS
BY
E. B. FORD
With 71 Photographs in Black and White
Taken By S. Beaufoy

(31) SPARE-PART SURGERY
BY
Donale Longmare
Edited and Illustrated By
M. Ross-Mac

(32) THE GOSPELS
AND THE JESUS OF HISTORY
BY
XAVIER LEON DUFOUR
Translated and Edited By John McHugh

the rule was formulated to: Recognize AND after BY in the substring, except when (a) AND is in the environment of BY ... WITH ... AND, (b) AND is in the environment of BY ... edited AND, BY ... AND edited.

With two exceptions, the recognition markers and environments specified in the rules required identical physical expression in the designated strings for recognition of the author to take place. The exceptions occurred with the markers BY and AND, which were found to be implied when they were not physically expressed in the author statement. For example:

(33) THE HIDDEN DIMENSION
Edward T. Hall

(34) Underground Storage of Fluids

By
Donald L. Katz
Keith H. Coats

(35) INTRODUCTION TO MORAL EDUCATION

John Wilson
Norman Williams
Barry Sugarman

where in (33), the marker BY is implied in the substring, that is, BY Edward T. Hall; where in (34), the marker AND is implied in the string, that is, By Donald L. Katz AND Keith H. Coats, and where in (35), both markers are found to be implied in the string, specifically BY John Wilson AND Norman Williams AND Barry Sugarman.

In some cases, the physical expression of these markers was indicated in the form of a slash:

(36) HOW TO GROW PLANTS ---- INDOORS AND OUTDOORS

By

Gerald Whitney/James Gooch/Doris Royce

or a comma:

(37) HUMAN BEHAVIOR

By
Richard Luchs, Carl Burnside, Eric Gross
and Lewis Findley

To accommodate the recognition of the author names involved, the rules were constructed to recognize these markers when implied in the substring, and to "write them in" by inserting them in brackets, [ ], in their appropriate position in the string. For any given substring without BY, the rules would recognize and write in BY at the head or commencement of the substring. For example, in the case of (33) above, the rules would identify the substring as: Edward T. Hall, and recognize and insert the implied marker BY at the head of the string: [BY] Edward T. Hall. For subsequent strings in the substring, the rules would recognize and write in AND at the end of each string except the terminal string in the absence of any other marker in the substring. For example, in (34), the rules would identify the substring as:

By
Donald L. Katz
Keith H. Coats

and in the absence of any other marker, recognize and insert AND in the designated position: By Donald L. Katz [AND] Keith H. Coats. In the case of slashes and commas, the rules would automatically substitute, that is, write in the marker in brackets as, for example, in (36): By Gerald Whitney [AND] James Gooch [AND] Doris Royce, and (37): By Richard Luchs [AND] Carl Burnside [AND] Eric Gross and Lewis Findley.

From this brief discussion, it is possible to detect some of the essential properties of the characterization. For one thing, it is clear that an order of application must be defined with respect to the use of these rules. The application of (A) _______ $\overset{\&}{\text{AND}}$ _________'S,S', for example, must apply before (A) _________'S,S', so that the two author names occurring in that pattern will be recognized. More specifically, before either of these rules, the rule for recognizing BY in the substring must first apply, since the applicability of the rules for the other markers is contingent upon whether or not this marker is present, expressed or implied. Following the rule for BY are the rules for recognizing markers in the initial or first string. These rules involve the markers 'S,S', OF, FROM, BETWEEN, and WITH, the sequence of their application being determined largely by their relation to each other. An example in point is the marker 'S,S' above.

Following this group of rules are the rules for recognizing markers in the substring. These rules involve the markers AND, WITH, and OF, with the sequence of their application likewise being determined largely by their relation to each other. For instance, the rule for AND calls for recognizing the marker in the substring with two exceptions. One exception is when the marker is in the environment BY ... WITH ... AND. This rule would precede the rule for the marker WITH, which is recognized except when it is in the environment of BY ... WITH ... BY, etc. If the latter rule were to apply first, a WITH occurring in this sequence would not be recognized, in which case the exception to recognizing AND would apply vacuously.

Secondly, it is essential that the rules be ordered in such a fashion that permits a procedure for their application. In the characterization developed here, the rules are obligatory; that is to say, each of the rules must be applied to the strings in question in order to produce the desired recognitions. For each title page, this involves moving through the rules only once, carrying along the recognitions produced with the application of each rule until reaching the last rule and the complete recognition.

Summary of the recognition component. -- The characterization of the recognition component for each corpus of entry rules, then, took the form described above. To ensure uniformity in application and for the purposes of comparison later, it was necessary to express the characterization in a standard form. The explanation that follows serves to summarize points of the characterization presented earlier, and to introduce the symbols and conventional notations used to express the characterization. For ease in interpretation, the symbols and notations are introduced in the context of their practical application as much as possible.

1. In the case of each corpus of main entry rules, the recognition component of the entry selection process is characterized as a set of rules, whose explicit function is to recognize in a completely objective, mechanical manner the presence of an author name or names on the title pages of English language books.

2. The title page data which serve as the basis for the construction and application of the rules consist of two statements commonly used in the determination of main entry, namely, the author statement and the title statement.

3. These two statements are identified by the rules as a sequence of two physical strings, of which the title statement is the initial or first string, and the author statement is the sub or second string.

each of the strings being composed of as many lines as there are in each statement.

4. The rules consider each of these strings to be in the form #S#, where the ends of each string are marked with a boundary marker #. The boundary marker represents the physical breaks on the title page that come before, between, and after the title and author statements and is a technique used by the rules to set apart and at the same time define the two strings. Taking the title page below,

Introduction to Geometry
By
G.A. Dickinson

the rules would dictate the insertion of boundary markers at the ends of each string, and define the string either as initial (first) or sub (second) string.

# Introduction To Geometry #	}	initial string
# by G. A. Dickinson #	}	substring

Both the construction and application of the rules developed in the study are dependent on the abstraction of the title page as noted above.

5. In the case of each corpus of main entry rules, the recognition component is characterized as a set of instructions of the form: Recognize such and such an element ..., where to recognize is to identify the exact shape and form of the specified element in the initial string, substring, or both strings as dictated, and where the element is any one of a

set of seven markers. BY, AND, WITH, OF, BETWEEN, FROM, 'S,S', given in the rules.

6. Since the markers are found to function in numerous contexts in addition to that of author name recognition, the rules place restrictions on their recognition by specifying the position and the environment in which they must occur on the title page. With respect to position, the rules take into account the physical separation of the author and title statements and the author recognition function of the markers in these statements. With respect to environment, the rules take into consideration certain words or elements in whose context the markers are found to function in the role of author indicating devices.

7. Where marker recognition is restricted to certain contexts only, the rules are of the form: Recognize such as such an element when it occurs in the environment of ...., where "in the environment of" is formally expressed in the rules as a slash, /, of the type:

$$\mathcal{X}/\mathcal{Y} \ldots \_\_\_\_$$

with $\mathcal{X}$ and $\mathcal{Y}$ being given marker elements and the rule reading: Recognize marker $\mathcal{X}$ in the environment of marker $\mathcal{Y}$, the dots, ..., representing anything or nothing coming after $\mathcal{Y}$, and the blank, _______, in the rule specifying the position of marker $\mathcal{X}$ in the environment. An example of a context-sensitive or -restricted rule might be: Recognize OF in the substring when OF/WITH ... _______, where the marker OF is recognized only when it comes after the marker WITH, the ____ indicating the position of the marker.

8. Often a number of environments apply to a given marker. In such cases the environments are listed vertically within braces, { }, the braces being used to collapse what would be several rules into a single rule. For example, Recognize OF in the initial string when

$$\text{OF} \;\Big/\; \left\{\begin{array}{l}\text{articles}\\ \text{conversations}\\ \text{dialogues}\\ \text{letters}\end{array}\right\} \underline{\qquad\qquad}$$

where the marker OF is recognized when it comes directly after "articles," when it comes directly after "conversations," when it comes directly after "dialogues," and so on, its position indicated with the ______. In a few cases, a number of the same environments apply to several markers. The formal expression and interpretation of the rule remain the same, with the markers positioned to the left of the slash, listed vertically and enclosed within braces.

$$\left\{\begin{array}{l}\text{OF}\\ \text{FROM}\\ \text{BETWEEN}\\ \text{WITH}\end{array}\right\} \Big/ \left\{\begin{array}{l}\text{articles}\\ \text{conversations}\\ \text{dialogues}\\ \text{letters}\end{array}\right\} \underline{\qquad\qquad}$$

9. In addition to a vertical listing within braces, some environmental elements are enclosed in parentheses, ( ). In such cases, the parentheses are used to indicate optional elements, elements that may or may not be present in the string. An example of an optional element that is added to the following rule is the indefinite article,

A: Recognize (A) ________'S,S' when it occurs without BY in the substring, which has the interpretation recognize the marker 'S,S' with the association of A when it occurs, recognize the marker 'S.S' when it occurs without the association of A.

10. With two exceptions, the markers and environments specified in the rules must have identical expression in the strings for recognition of the author's name to take place. The exceptions occur with the markers AND and BY, which in many cases are implied although not physically expressed. The rules for these markers take cognizance of this condition and are stated in the form: Recognize BY, expressed or implied, in the substring, etc. The marker BY is said to be implied when a substring is present and there is no physical expression of the marker in it. In these cases, the rule dictates that the marker be inserted in brackets, [ ], at the heading of the substring. The marker AND is said to be implied when there is a substring and no physical expression of this and any other marker in it. In these cases, the rule dictates that the marker be inserted in brackets at the end of each string except the terminal string in the substring.

11. Finally, with respect to their application, the rules are strictly ordered. Application starts at rule 1 and continues with the application of successive rules 2, 3, and so on through to the last rule, this procedure being followed for each title page. With the application of each rule, the appropriate strings are examined for the marker and environments in question. If the marker and environments are identical to those specified in the rule, the recognition is copied down. With the application of each succeeding rule the recognitions copied

down from the previous rules are carried forward and noted in the appropriate strings and in the order in which they appear in their strings. Two lines for the two strings (initial) and (sub), are provided for the application of each rule. With the application of the last rule, author name recognition terminates, the complete recognition being the last one copied down under the rules.

A presentation of the recognition rules developed for each corpus of main entry rules is given in the next section beginning on page 162.

2. The selection component. ---- On the basis of the recognition component, the selection component in both the AA and ALA cataloging codes functions to assign the main entry to the person or corporate body considered to be chiefly responsible for the intellectual content of a work. Using the results, that is, the recognitions produced by the recognition component, the selection component in both codes operates to choose the appropriate main entry, an author entry when intellectual responsibility can be assigned, or a title entry when authorship or authorship responsibility is lacking. The way in which each selection component is seen to accomplish its function is through the use of certain principles and rules devised especially for this purpose. For each component in the two codes, the principles and rules of entry are formulated as sets of verbal instructions of the form: Enter a work ...

These instructions specify which main entry is permissible and which one of several items is permitted in the entry, taking into consideration the various patterns of authorship and the various types of materials in which these patterns are displayed. The application of these instructions is, of course, dependent upon the cataloger, whose training and native capacities to understand and interpret their verbal import are essential to the component. To construct a characterization of the selection component, it will be necessary to devise some way to represent both the structure and content of these instructions and to account somehow for the cataloger's ability to apply them.

It is apparent that characterization of this component presumes first the characterization of the recognition component. For the characterization to specify an entry and the item that is to occupy that entry, it must have at its disposal the output produced by the recognition component, namely, the strings consisting of author names and markers and lexical units associated with them on the title page. Similarly, in the case of the cataloger, selection of entry presumes the cataloger's examination of the title page, the items of information on it, their typographical prominence, physical arrangement, etc. It is this output of the recognition component then or, in the case of the cataloger, the results of his examination of the title page data, etc., that provides the selection component with the input needed for making the appropriate main entry choice.

In both cataloging codes, the principles and rules of entry are constructed in terms of the various patterns that authorship assumes — single authorship, joint or shared authorship, unknown authorship, etc.

It can likewise be expected that a characterization of this component will be constructed in a similar manner, using the various patterns of authorship that comprise the content of the verbal principles and rules.

A case in point might be ALA rule 3 for joint authorship and AA rule 3 for shared authorship, where the two rules have as their content a pattern of two or more authors. In both rules, the pattern is defined to apply to author statements, where the names are found to be conjoined or linked by ANDs, WITHs, and OFs; for example: By WILL AND Ariel Durant, By Dai Rees WITH John Ballentine, By Michael C. Hall with the assistance OF Dorothy S. Kinoshita, etc. (the conjunctions, etc. capitalized for emphasis). To represent these two rules in terms of the authorship pattern they specify, a characterization would utilize the structures of the author statements to which the rules applied, since it would not, of course, comprehend or interpret the verbal import of the rules themselves.

For instance, in the case of these two rules, the structures BY x AND y, BY x WITH y, BY x OF y (where the x and y represent the author names in the structures) would be used as a characterization to represent the authorship pattern comprising these rules. Each of these structures would be incorporated in the characterization as a representation of these particular rules. The structure BY x and y, for instance, would be incorporated as one representation of the authorship pattern in the rules; BY x WITH y would be incorporated as another, and so on.

Two difficulties that can be expected to arise in connection with this particular characterization are: (1) the number of structures which may be found to represent a given rule, and (2) the number of identical structures which may be found to represent different rules. With respect to (1), it has already been seen from the examples above that there are several structures representing the rule for joint authorship in the ALA code and shared authorship in the AA code. With respect to (2), it is sufficient to take a simple structure such as BY χ to illustrate this predicament. This structure is one of several that characterizes the rules for single authorship in the two codes, as for example: By Victor B. Scheffer. It is likewise a structure that is found to characterize the rule for unknown authorship in the AA code and anonymous authorship in the ALA code, as for example: By A Servant. By Eight Leading Assessing Officers, etc. These difficulties are seen to pose problems specifically with respect to the simplicity of the simulation models and their operational capacity. In short, the characterization must be constructed with an economy of structures representing the verbal principles and rules. In the case of identical structures representing different rules, it must have some method of examining the structures that will ensure the selection of the main entry specified by those particular verbal rules.

In this connection, it is expected that a characterization of the selection component will embody those processes that relate to the cataloger's use of the verbal principles and rules and performance of the task. The mental processes that a cataloger uses in applying the rules and determining an entry choice are undoubtedly complex. However, it appears that, given the recognition of authorship, the cataloger adopts a system or some routine, whereby he applies a particular rule or rules in some defined sequence, reflecting perhaps the order of the rules in the code or, more likely, the order and arrangement of the authors on the title page. With the application of a rule, it is reasonable to suppose that the cataloger, for example, examines the pattern of authorship on the page to see if it agrees with the pattern contained in the verbal rule. For instance, if the rule for either joint or shared authorship were being applied to a work, the cataloger would examine the title page to determine if the pattern of authorship was one, where two or more author names were conjoined in accordance with the content of the rule. If the pattern were found to agree with that contained in the rule, the cataloger would terminate the selection process with the entry choice specified by the rule in question. In the event that the pattern did not satisfy the rule, the cataloger would be expected to continue to apply successive verbal rules until an agreement and an entry decision were reached.

If this system were adopted by a characterization, it would have as the verbal principles and rules of entry the structures of the author statements to which the rules applied. Given this characterization, the application of the structures to the recognitions produced by the recognition component would utilize a system similar to that inferred to the cataloger. For instance, given a recognition, a structure characterizing a particular rule would be applied. In the case of an agreement, that is to say, where the recognition would be found to "match the structure of the rule, the selection of a main entry under that particular rule would be considered to occur. To specify the main entry dictated by the rule, the recognition, for example, would be examined for the presence of certain physical characteristics -- typographical prominence of the author names, indefinite articles introducing the author names, etc. that are used by the verbal rules as indications of the presence or absence of authorship responsibility. On the basis of this examination, an entry decision would either be specified or deferred for the application of other structures representing that or other rules.

Much more is needed in the way of defining the system. In particular, some sort of heuristic procedure must be supplied which will enable the characterization to examine a recognition for each of the physical characteristics mentioned, and to produce the entry required for each alternative. Similarly, a heuristic procedure must be supplied, which will enable the characterization to apply the structures to a recognition in an ordered sequence until one is found to apply. While these procedures remain to be developed, it nonetheless seems reasonable to assume that a system of the type mentioned underlies the cataloger's use of the rules of entry and his performance of the task.

The characterization of the selection component developed in the study was based on the system described above. In the following paragraphs the construction of the characterization is described in greater detail.

For each corpus of main entry rules, the selection component was characterized as a set of rules for determining authorship responsibility. In the case of each corpus, the rules were constructed in such a way that when applied to the author name recognition produced by the recognition component an appropriate main entry is assigned, either an author entry when intellectual responsibility can be determined, or a title entry when authorship or authorship responsibility is lacking.

These selection rules were constructed from the verbal main entry rules in each corpus. Each verbal rule was analyzed in terms of the pattern of authorship it contained — single authorship, joint or shared authorship, unknown authorship etc. and the entry it specified — author where the author is an editor, first name in the author statement, etc., or title where there are four or more authors in the author statement no principal author, etc. In essence, this necessitated reducing the verbal construct of each rule to its fundamental structure, a task which was accomplished earlier in the study when structural descriptions were assigned to the rules for testing the hypothesis. Although these structural descriptions, given in Tables 7 and 8 in Chapter III, could not be used as the rules in the characterization, they provided certain features that could be used in constructing a set of selection rules.

For instance, in the structural description of AA 8A below:

Work of artist and author → A(where $A = \begin{Bmatrix} \text{artist} \\ \text{author} \end{Bmatrix}$ = first personal name)

$$\begin{bmatrix} \text{+artist} \\ \text{+author} \\ \text{+collaboration of artist and author} \\ \text{+identification of artist and author} \\ \quad \text{on title page} \end{bmatrix}$$

Unless second personal name has typographical or lexical prominence

Then $A = \begin{Bmatrix} \text{artist} \\ \text{author} \end{Bmatrix}$ = second personal name

those features, such as artist and author, are explicit, that is to say, directly observable as to their presence on the title page, and could be used by a characterization to represent the rule. Indeed, the identification of the artist and author on the title page is a feature essential to the application of the rule. In contrast, a feature such as the collaboration of the artist and author might be considered implicit. With the exception of an author statement indicating the name of the artist "with the collaboration of" the writer, a characterization, of course, would have no way of confirming from the title page whether this feature were present or absent. Although this feature is similarly essential to the application of the rule, it is nonetheless possible that the verbal import of the rule could be adequately characterized by using those features that are directly observable on the title page. In short, it may be that successful selection for the main entry can be achieved despite the inability of the characterization to always represent the feature.

For example, taking AA 8A above, the features of artist and author in the rule are directly observed in author statements such as:

(38)

BY

Richard Hohn

CARTOONS BY
T. K. Whitehead

(39)

Photographs By
Lois Wright
Text By
Benjamin Caughey

where the construct of the statements is: BY $\mathcal{X}$ BY $\mathcal{Y}$, the $\mathcal{X}$ and $\mathcal{Y}$ being observed in the structure as either an author and artist name, or an artist and author name. By using this structure of the author statements, the verbal rule with its specification of entry under the first name in the author statement could be characterized as a rule of the form: BY $\mathcal{X}$ BY $\mathcal{Y}$ $\rightarrow$ A, where A = $\mathcal{X}$, with the general interpretation that an author statement having this structure goes to or is entered under author, where the author to occupy the entry is $\mathcal{X}$, the first name in the author statement.

At the first level of development, the verbal rules of entry in each corpus were characterized in the manner and form described above. In the case of each rule, the features contained in the rule and essential to its application were examined for their occurrence and use in the author statement. Where the features were directly observable, the structure of the author statement was noted. The structures to which a rule was applicable supplied the characterization with a means of representing the verbal import of the rule. Specifically, this involved using the structures to create rules which incorporated the features and specifications of entry contained in the verbal rule.

A difficulty in connection with using this approach is that there are, as previously mentioned, any number and variety of author statements to which a single verbal rule will apply. On the one hand, this results in a multiplicity of rules to characterize the verbal rule. On the other hand, it results in an incomplete characterization of the verbal rule. For instance, in the case of AA 8A above, it was found that in addition to the structure BY $\mathcal{X}$ BY $\mathcal{Y}$, the rule was also applicable to author statements such as the following:

(40)

DRAWINGS BY
Thomas Finn

CAPTIONS BY

Richard McElroy and Lloyd Power

(41)

Text By
Jesse Tate and Florence Burns

Paintings By
Paul R. Bozzini

(42)

BY
JOHN P. GOULET
AUTHOR

AND
GERALD HEALY
PHOTOGRAPHER

where the features of artist and author are found in the constructs, BY $\mathcal{X}$ BY $\mathcal{Y}$ AND $\mathcal{Z}$, BY $\mathcal{X}$ AND $\mathcal{Y}$ BY $\mathcal{Z}$, and BY $\mathcal{X}$ AND $\mathcal{Y}$. The applicability of the rule to these structures results in the creation of several rules to characterize the verbal rule. At the same time, it is apparent that the characterization of the rule is incomplete, since there is obviously no way of knowing or anticipating all the structures to which the rule will be applicable. This means that the rules in the characterization are restricted to apply only to those author statements whose constructs are already known.

In order to resolve these difficulties, it was necessary to rephrase the large number of special rules in the characterization in the form of fewer, more general-type instructions. In the case of each verbal rule, it was possible to express the rules in the characterization in one or two rules. For example given the four rules constructed for AA 8A,

BY $\mathcal{X}$ BY $\mathcal{Y}$ → A, Where A = $\mathcal{X}$
BY $\mathcal{X}$ BY $\mathcal{Y}$ AND $\mathcal{W}$ → A, Where A = $\mathcal{X}$
BY $\mathcal{X}$ AND $\mathcal{Y}$ BY $\mathcal{W}$ → A, Where A = $\mathcal{X}$
BY $\mathcal{X}$ AND $\mathcal{Y}$ → A, Where A = $\mathcal{X}$

it was found that the following rule,

BY $\mathcal{X}Z$ → A
Where A = $\mathcal{X}$
And Z = anytning or nothing

could be posited in place of those rules. This rule had the interpretation that an author statement having any one of these structures goes to or is entered under author, where the author to occupy the entry is $\mathcal{X}$, and where the Z in the structure is anything (any other author name or names in the statement) or nothing (no other author name in the statement). With the construction of this rule, which incorporates the features and entry specification contained in the four separate rules, the characterization was greatly simplified. At the same time, the rule amounts, in effect, to a generalization about all the structures to which the verbal rule may ever be found to apply. That is to say, the rule incorporates not only those structures in the rules above, but all possible structures to which the verbal rule would be applicable.

As in the case of the recognition rules, the selection rules were constructed to take into consideration the position of the author names in the statements, and the environment in which they are found to occur. Since the verbal main entry rules in the two cataloging codes are context-sensitive or -restricted, that is to say, an entry is specified only in certain contexts, it could be expected that this would also be true of the rules in the characterization. A typical example is AA 9A, where a publication consisting of a work or group of works of an author, which is accompanied by biographical or critical material by another person, is entered under the latter provided he is represented as an author and not as an editor. The illustration provided in the rule is:

(43) Life and letters of Mrs Jason Lee ... by Theresa Gay

where the main entry is under Gay. In the characterization, this rule is expressed in a selection rule of the form:

$$\text{Life and} \begin{Bmatrix} \text{letters} \\ \text{poems} \\ \ldots \end{Bmatrix} \text{ of } xZ \text{ BY } yZ \rightarrow A$$

where A = y

And Z = anything or nothing but not EDITED
before or after BY

This predicament arises in AA 2, for example, where it is found that the verbal rule shared, in addition to other structures, some of those which characterized AA 1A. An example is:

(44)
Successful Selling
Of
Crafts
By
The Leading Craftsmen

(45)
CERTITUDES
By
Sister M. Leanore

where the structure of the two author statements is BY $\mathcal{X}$. In the case of (44), AA 2C applies to assign a title main entry. In (45), AA 1A applies to assign an author main entry. In the corpus of ALA rules, this predicament arose in a number of rules, among them ALA 5A(1a) and ALA 2. For example:

(46)
Famous Irish Fairy Tales
Introduction By
CONOR REDDIN

(47)
THE SUN ALSO RISES
by
Ernest Hemingway

where both author statements have the structure: BY $\mathcal{X}$. However, in (46), ALA 5A(1a) applies to assign a title main entry while in (47), ALA 2 applies to assign an author main entry.

In these and other cases, it was necessary to build exceptions into existing rules in the characterization. These exceptions were based on certain physical characteristics that were observed in the data and, in most cases, were used by the verbal rules to accomplish the desired main entry. For instance, in (44) and (45), where AA 2C and AA 1A are applicable to author statements with the construct: By $\mathcal{X}$, it was possible to take the general rule:

BY $\mathcal{X}Z \rightarrow A$
Where $A = \mathcal{X}$
And Z = anything or nothing

and build in the following exception:

Unless $\mathcal{X}$ = the (definite article)
Then $\rightarrow T$
Where T = title of monograph

With the exception, namely, that when an author name is introduced with a definite article, entry is under title, the rule could apply to recognitions such as: By Sister M. Leanore in (45) to successfully accomplish entry under author and, at the same time, to recognitions such as: By the Leading Craftsmen in (44) to successfully accomplish entry under title.

In (46) and (47), where ALA 5A(1a) and ALA 2 are applicable to author statements with the same construct: BY $\mathcal{X}$, it was similarly possible to take an identical rule.

BY $\mathcal{X}Z \rightarrow A$
Where A = X
And Z = anything or nothing

and build in the exception.

Unless $\mathcal{X}$ is in the environment of Introduction BY
Then → T
Where T = title of monograph

With this exception, the rule was able successfully to accomplish entry under author in cases such as (47) where the recognition is: by Ernest Hemingway, and entry under title in cases such as (46) where the recognition is: Introduction By Conor Reddin. Since both exceptions applied to the two identical rules, these were incorporated in each statement of the rule as follows:

BY $\mathcal{X}Z \rightarrow A$
Where A = $\mathcal{X}$
And Z = anything or nothing

$$\text{Unless } \mathcal{X} = \left\{ \begin{array}{l} \text{the} \\ \qquad [\pm\text{definite}] \\ \text{Introduction BY } \_\_\_ \end{array} \right\}$$

Then → T
Where T = title of monograph

With the vertical listing of the exceptions in braces, { }, it was possible to collapse the multiple exception statements into one unified statement, and thereby simplify the characterization.

In addition to wording, certain of the verbal rules in each corpus state or imply the use typography to assist in the determination of authorship responsibility. For example, rule 3A in the AA code calls for entering a work of shared responsibility under

> the person or corporate body, if any to whom principal authorship is attributed, e.g. by wording or typography ...

In the ALA code, rule 5A(1a), for example, calls for entering collections under title

> if the work of the editor or editing body seems to be but slight and their names do not appear prominently ...

This physical characteristic was similarly employed by the rules in the characterization to ascertain authorship responsibility. As in the case of the wording, it was built into the rules as an exception of the form:

BY $\mathcal{X}Z \rightarrow A$
Where $A = \mathcal{X}$
And Z = anything or nothing
Unless Z is in larger type
Then $A = Z$

Given a title page display such as:

Science Experiments
For Children
By
John Gavin With MARY SCHATZ

where the recognition would be: By John Gavin With MARY SCHATZ, this rule would function to assign the main entry to Schatz, in accordance with AA 3A above.

The applicability of different verbal rules to author statements with identical constructs necessitated the incorporation of exceptions into existing rules in the characterization. At the same time, it permitted a further reduction in the number of rules needed to characterize each set of verbal rules. While it had been possible at first to reduce the large number of separate special rules in the characterization by constructing one or two more general rules for each verbal rule, it was now possible to collapse over half of the rules in the characterization of each corpus by constructing still more general rules. The manner and form of this reduction are illustrated in the examples given above for AA 1A and AA 2C and ALA 2 and ALA 5A(1a), where the separate rules for each pair of code rules were collapsed into one rule by building in the exceptions noted. Another typical example is the reduction of rules characterizing ALA 5C(4) and ALA 23C. A general rule characterizing ALA 5C(4) was of the form:

$$\left\{\begin{matrix}\text{essays}\\ \text{tales}\\ \dots\end{matrix}\right\} \quad \text{FROM } \mathcal{Z}Z \rightarrow A$$

$$\text{Where } A = \mathcal{Z} = \text{serial} \quad \left\{\begin{matrix}\text{Highlights}\\ \text{Life}\\ \dots\end{matrix}\right\}$$

$$\text{And } Z = \text{anything or nothing}$$

while one characterizing ALA 23C was of the form

$$\left\{\begin{matrix}\text{essays}\\ \text{tales}\\ \ldots\end{matrix}\right\}\ \text{FROM}\ \mathcal{Z}Z \rightarrow A$$

Where $A = \mathcal{Z}$ = personal name
And Z = anything or nothing

An examination of the two rules indicated that one rule could be constructed to perform the functions of both rules, specifically:

$$\left\{\begin{matrix}\text{essays}\\ \text{tales}\\ \ldots\end{matrix}\right\}\ \text{FROM}\ \mathcal{Z}Z \rightarrow A$$

Where $A = \mathcal{Z}$
And Z = anything or nothing

With the construction of this rule, title page displays such as:

(48) ESSAYS
From LIFE MAGAZINE

(49) MACABRE TALES
FROM
Edgar Allan Poe
Selected By
P. W. Doyle

whose recognitions would otherwise require the application of the two rules, could be successfully handled.

Finally, an obvious reduction in the number of rules in the characterization was realized when rules were found to be identical. In the case of each corpus, the rules affected had been characterized with the rule:

$\text{BY } \mathcal{X}Z \rightarrow A$
Where $A = \mathcal{X}$
AND Z = anything or nothing

From this discussion it is apparent that some order of application must be imposed with respect to the use of the rules. A rule such as:

$\mathcal{X} \overset{\&}{\text{AND}} \mathcal{Y} \text{ 'S,S' } Z \text{ BY } \mathcal{W}Z \rightarrow A$
Where $A = \mathcal{X}$
And Z = anything or nothing

must precede:

$\mathcal{X}\text{'S,S' } Z \rightarrow A$
Where $A = \mathcal{X}$
And Z - anything or nothing

for instance, in order to avoid entry under $\mathcal{Y}$. This rule would apply to recognitions of the type: Pratt and MacKenzie's (Law of Highways 21st Edition) By Harold Parrish and Lord de Mauley, where under AA 3 or ALA 20, main entry would be assigned to the original author.

In turn, the rule,

$\mathcal{X}$'S,S' Z → A
Where A = $\mathcal{X}$
And Z = anything or nothing

must precede a rule like,

BY $\mathcal{X}$ AND_1 $\mathcal{Y}_1$ AND_2 $\mathcal{Y}_2$ AND_3 $\mathcal{Y}_3$ Z → T
Where T = title of monograph
And Z = anything or nothing

to permit successful selection of entry for recognitions such as: Twains's (Humorous Anecdotes Selected) By Ben Paish [AND] Ruth Larson [AND] M.G. Legge [AND] Rachel Edmonds, where under AA 1a or ALA 23C, main entry would be assigned to the author of the selections.

Similarly,

BY $\mathcal{X}$ AND_1 $\mathcal{Y}_1$ AND_2 $\mathcal{Y}_2$ AND_3 $\mathcal{Y}_3$Z → T
Where T = title of monograph
And Z = anything or nothing

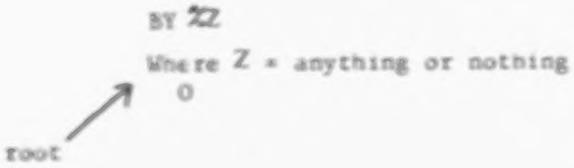
must precede the rule:

BY $\mathcal{X}$Z → A
Where A = $\mathcal{X}$
And Z = anything or nothing

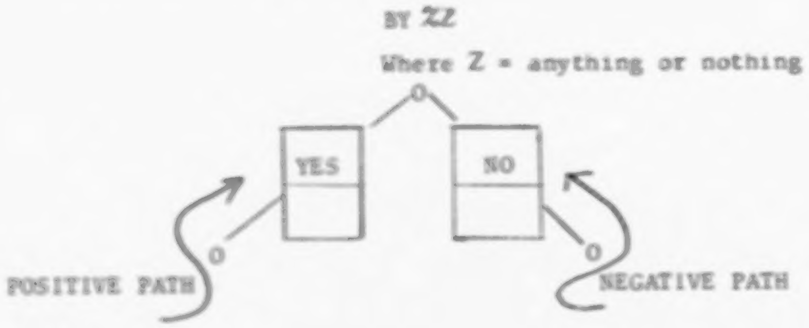
to allow recognitions such as: BY RICHARD LAKOFF [AND] ALICE MANNING [AND] GREGORY L. BUSH [AND] DELBERT HYNES to be entered under title in accordance with AA 3B2 or ALA 4A(2). In short, the order of application of the rules in the characterization was necessarily determined by the degree of restrictiveness in the rules. The first rule in the characterization, for instance, was the most restrictive, applying to the fewest or least number of recognitions. This rule was followed by successive rules of an increasingly less restrictive nature, applying to an increasingly larger number of recognitions. The last rules in the characterization for the most part were the least restrictive or most general rules, applying to the largest number of recognitions. With a reverse in the ordering, the most general rules would have the effect of applying to all recognitions, discounting any real need or use for the other rules in the characterization.

In addition to ordering the application of the rules, it is also apparent that a characterization must be supplied with some method of accepting a recognition and examining it for those physical characteristics that assist in the determination of entry. In other words, the characterization must have, in addition to the selection rules, a capacity to perform logical operations of various sorts. It must, for instance, be able to take a recognition and first compare its structure with that contained in a selection rule. On the basis of a match or an agreement, it must then be able to test the recognition on each one of the physical characteristics used by the rules — typography, wording, etc. In terms of the presence of a characteristic, it must assign an entry and specify the item to occupy the entry, or defer entry selection for the examination of other alternatives, or the application of still other selection rules.

To perform the operations described, it was necessary to supply the characterization with a heuristic component, namely, a set of procedures or techniques which would permit a sequential application of the rules, and a sequential analysis of the recognitions accepted by the rules. In the characterization developed in the study, these procedures involved a graphic representation of the selection rules in the form of abstract trees and the use of a binary decision process to facilitate their application. A description and an illustration of the tree graphs used in the study are given in Chapter II (see p. 15). For each selection rule in the characterization, a tree was constructed, which had at the root or highest node the construct of the rule in the form:


BY ZZ
Where Z = anything or nothing
O
root

From the root, two lines or branches extended downward to two lower nodes. These branches represented positive and negative paths, indentified as Yes-No responses of the form:


BY ZZ
Where Z = anything or nothing
O
YES
NO
O
O
POSITIVE PATH
NEGATIVE PATH

Given a recognition produced under the recognition rules, the selection rule would apply to compare the structure of the recognition with the construct of the rule contained at the root. If the structure of the recognition should match the construct of the selection rule, it would be passed to the node to the lower left. If the recognition should not match the construct at the root, the recognition would be passed to the node to the lower right. For instance, given a recognition such as: By John V. Smith and Dodie Baker, application of the selection rule: BY %Z → A would indicate a match or agreement between the structure of the recognition and the construct of the selection rule.

Construct of the rule: BY % Z

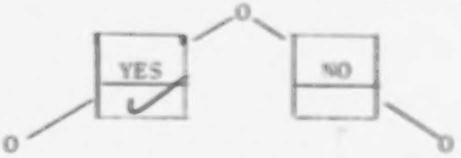
Where Z = anything or nothing

Recognition: By John V. Smith And Dodie Baker

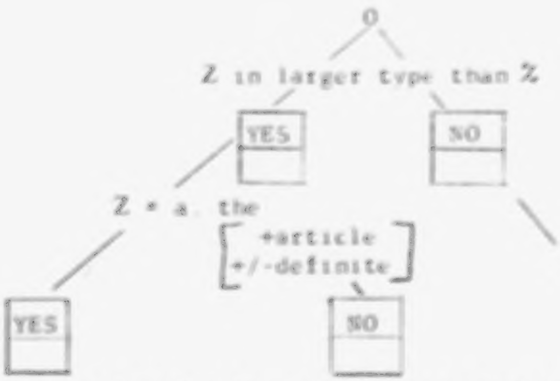
This match would pass the recognition down the left or positive path of the tree to a lower node:

BY %Z

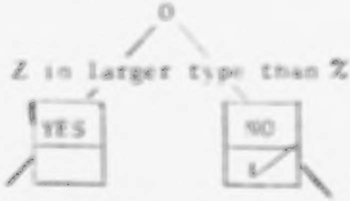
Where Z = anything or nothing



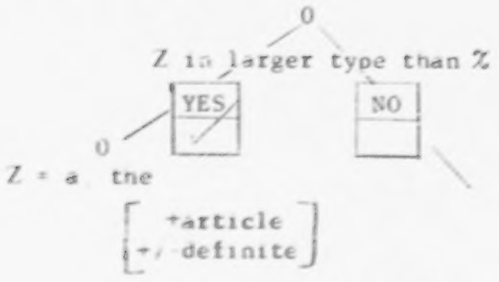
Associated with the lower nodes in the trees were specifications that a recognition must satisfy, positively or negatively. These specifications were the physical characteristics of typography and wording for which the recognitions were examined in a sequential analysis of the type:


0
Z in larger type than Ƶ
YES
NO
Z = a, the
+article
+/-definite
YES
NO

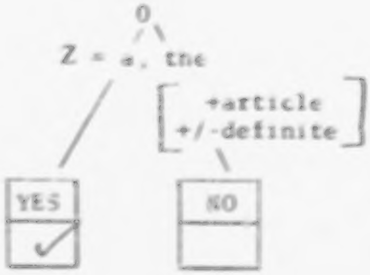
and so on. At each node, the recognition produced under the recognition rules would be examined for the specification. For the recognition: By John V. Smith And Dodie Baker, examination of the typography would indicate that Z (And Dodie Baker) is not in larger type than Ƶ (John V. Smith). This recognition would be passed down the right or negative path of the tree in the form of:


0
Z in larger type than Ƶ
YES
NO

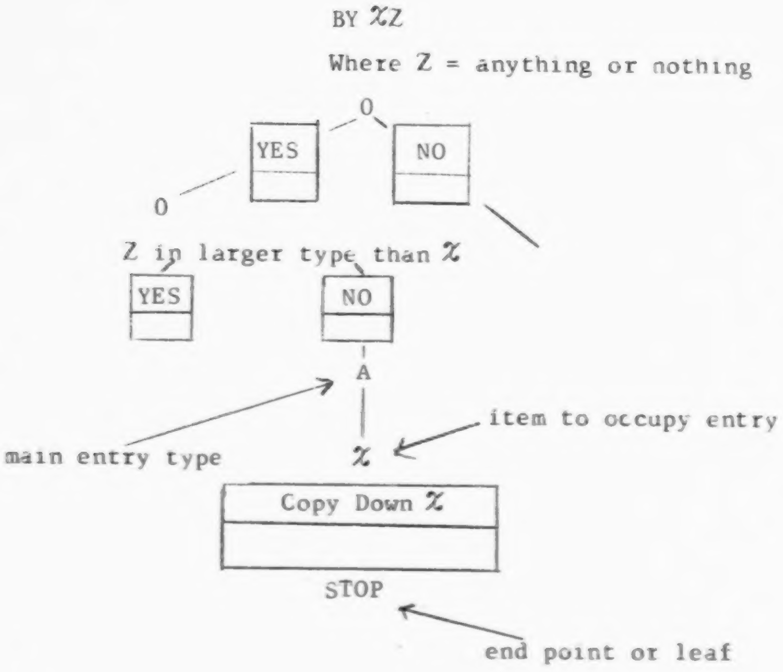
For a recognition of the same structure: By Private Kent Salar And A NAMELESS BUDDY, examination of the typography would indicate a match, that Z (And A NAMELESS BUDDY) is in larger type. Consequently, the recognition would be passed down the positive path.


0
Z is larger type than Z
YES
✓
NO
0
Z = a, the
+article
+/-definite

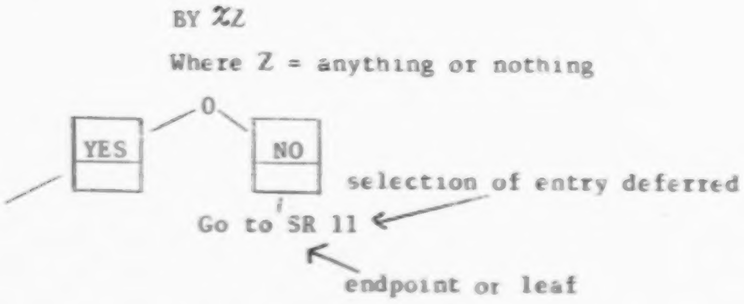
to the lower node in the illustration, where the specification is: Z = a, the, an indefinite or definite article. Since examination of Z indicates a match, namely, that Z (And A NAMELESS BUDDY) is introduced with the indefinite article, a, the recognition would continue down the positive path:


0
Z = a, the
+article
+/-definite
YES
✓
NO

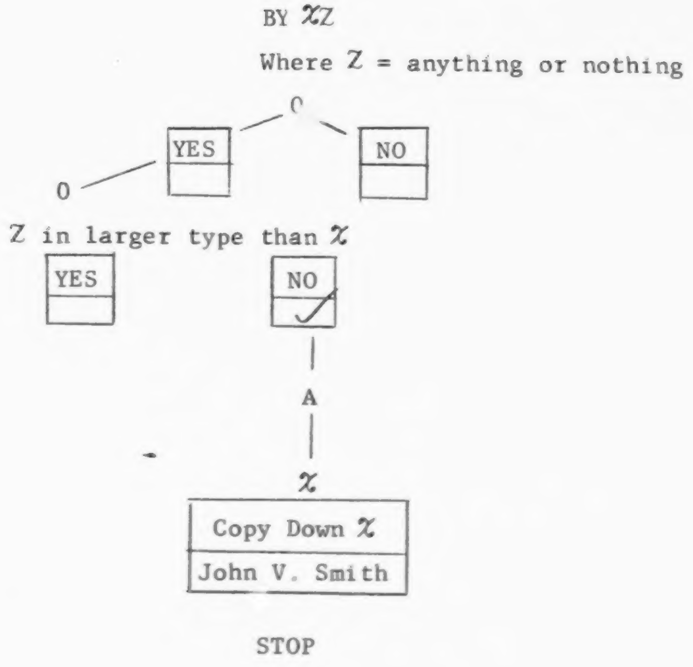
until an endpoint or leaf was reached. This endpoint or leaf was a node with no nodes below it. In the case of each selection rule in the characterization, the endpoints in the tree specified the entry for the rule, A (author) or T (title), and the item to occupy the entry, an author name or title, for example.


BY 𝒳Z
Where Z = anything or nothing
0
YES
NO
0
Z in larger type than 𝒳
YES
NO
A
item to occupy entry
main entry type
𝒳
Copy Down 𝒳
STOP
end point or leaf

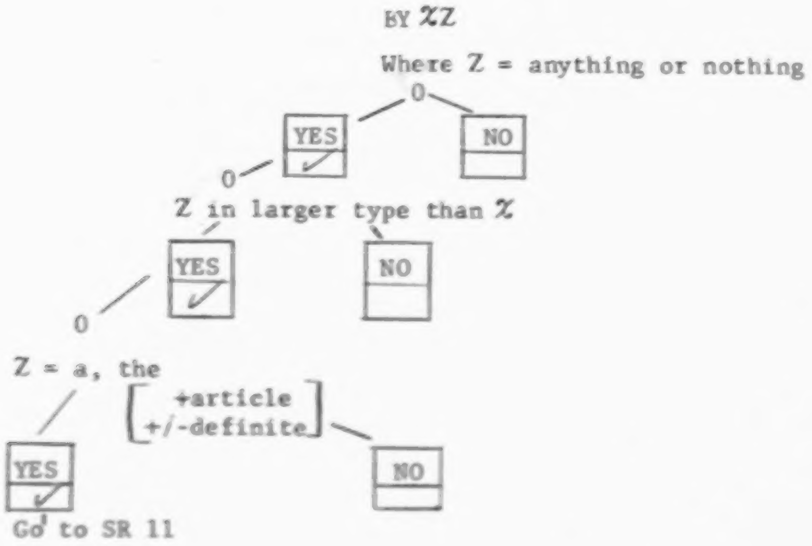
or deferred selection of entry for the application of the next rule, as for example:


BY 𝒳Z
Where Z = anything or nothing
0
YES
NO
selection of entry deferred
Go to SR 11
endpoint or leaf

For the recognition By John V. Smith And Dodie Baker, the endpoint in the selection rule would specify an A (author) entry, where the author is $\mathcal{X}$, John V. Smith:


BY 𝒳Z
Where Z = anything or nothing
0
YES
NO
0
Z in larger type than 𝒳
YES
NO
✓
A
𝒳
Copy Down 𝒳
John V. Smith
STOP

In the case of the recognition: By Private Kent Salar And A NAMELESS BUDDY, the endpoint in the rule would specify the application of the next selection rule:


BY 𝒳Z
Where Z = anything or nothing
0
YES
✓
NO
0
Z in larger type than 𝒳
YES
✓
NO
0
Z = a, the
+article
+/-definite
YES
✓
NO
Go to SR 11

As in the case of the recognition rules, application of the selection rules is obligatory, that is to say, the rules must be applied to a recognition in their listed order. In contrast to the recognition rules, however, the selection rules in the characterization apply only until the specifications of entry are satisfied. Upon reaching an endpoint with the designation of entry, the selection process terminates.

Summary of the selection component. — For each corpus of main entry rules, the characterization of the selection component was developed in the manner and form described above. As in the case of the recognition component, to ensure uniformity in application and for the purposes of comparison later, it was necessary to express the characterization of the component in a standard form. The explanation that follows serves to summarize points of the characterization presented earlier, and to introduce the symbols and notations used to express the characterization. For ease in interpretation, the symbols and notations are introduced in the context of their practical application as much as possible.

1. In the case of each corpus of main entry rules, the selection component of the entry process is characterized as a set of rules, whose explicit function is to identify author responsibility in a completely objective, mechanical manner. On the basis of the author name recognition produced under the application of the recognition rules, the selection rules operate to choose the appropriate main entry, either an author entry when authorship responsibility can be assigned,

or a title entry, when authorship or authorship responsibility is lacking.

2. The function of the selection rules and the basis for their construction are the verbal principles and rules of entry in the AA and ALA cataloging codes. For each corpus of verbal principles and rules, there is a set of selection rules constructed from those features of the verbal rules that are directly observable in the title page data, namely, the author and title statements. Using the features of the verbal rules, specifically, the author name or names, and the structures of the title and author statements in which they are found, the selection rules are formulated as instructions of the form:

$$\text{BY } \mathcal{Z} \rightarrow \text{A}$$
$$\text{Where A} = \mathcal{Z}$$

with the interpretation that an author name $\mathcal{Z}$ in a recognition where the structure of the author statement is BY ... goes to (→) or is entered under author, where the author assigned the entry is $\mathcal{Z}$.

3. Since there is a multiplicity of author statements to which a verbal rule may be applicable, the selection rules utilize the symbols Z and $\mathcal{Q}$ to represent the complexity and variety of such statements and the features of the verbal rules that may be found in them. These symbols are incorporated in the rules in the form:

$$\text{BY } \mathcal{Z}\text{Z} \rightarrow \text{A}$$
$$\text{Where A} = \mathcal{Z}$$
$$\text{And Z} = \text{anything or nothing}$$

with the interpretation that an author name $\mathcal{Z}$ by itself or with Z, any other author name or names, in a recognition with this structure is entered under author, where the author of the entry is $\mathcal{Z}$. This particular rule, for instance, would be applicable to recognitions such as: By John V. Smith And Dodie Baker, where $\mathcal{Z}$ in the recognition is Smith and Z is And Dodie Baker.

4. In a number of verbal rules, the features of the rules are observed in author statements having the same structure. To facilitate entry selection for those cases where the verbal rules are found to designate different entries for author statements with the same structure, the selection rules place restrictions on the position of the features in the statements and the environment in which they occur.

5. Where a main entry is assigned only when the conditions of position and environment are met, the selection rules are of the form: When an author name or names are in this position and environment in a recognition ... the entry is ..., which is formally expressed with rules of the type:

$$\mathcal{Z} \Big/ \begin{array}{l} \text{drawings BY } \underline{\hspace{2.5cm}} \rightarrow T \\ \text{Where } T = \text{title of monograph} \end{array}$$

where the main entry is designated as title, when an author name $\mathcal{Z}$ comes after drawings BY; the slash in the rule, / , indicating the environment in which $\mathcal{Z}$ must occur, and the blank, __________, indicating the position of the author name in the environment. A rule such as this would apply to a recognition with the statement: Drawings By Lettie Wolfe.

6. Often the same entry is designated for a number of environments. In such cases, the environments are listed vertically within braces, { }. As in the case of the recognition rules, the braces are used to collapse what would be several rules into a single rule. These rules are of the type:

$$\mathcal{X} \Bigg/ \left\{\begin{array}{l}\text{drawings}\\ \text{foreword}\\ \text{illustrat(ed) (ions)}\\ \text{introduc(ed) (tion)}\end{array}\right\} \text{BY} \;\_\_\_\_\_\_\_\_\_ \rightarrow T$$

Where T = title of monograph

where the main entry is designated as title, when an author name $\mathcal{X}$ comes after drawings BY, or foreword BY, or illustrated BY, or illustrations BY, etc. A rule of this kind would be applicable to recognitions having the statements: Illustrated By Sean Harris, Introduction By Patrick Drinan, etc.

7. For each set of selection rules, there is a heuristic component, namely, a set of procedures or techniques that enables the rules to accept a recognition and to examine it for certain physical characteristics that assist in the determination of main entry. In the characterization developed in the study, the selection rules are represented graphically in the form of trees, and accomplish entry selection through the use of a binary (YES-NO) decision process. At the root or highest node of each tree is the construct of the rule: BY $\mathcal{X}Z$, BY $\mathcal{X}$ AND$_1$ $\mathcal{Y}_1$

$AND_2\ \mathcal{Y}_2\ AND_3\ \mathcal{Y}_3\ Z$, etc. From the root, two lines or branches representing positive and negative paths extend downward to two lower nodes. If a recognition produced under the recognition rules should match the construct of the selection rule, the recognition would be passed down the positive path to a lower node. If the recognition should not match the construct, it would be passed down the negative path to a lower node. Associated with each lower node is a specification that the recognition must satisfy, positively or negatively. In the characterization, these specifications are the physical characteristics of typography and wording which assist in the determination of the main entry. If a recognition should satisfy a specification, that Z is in larger type than $\mathcal{X}$, for example, it would pass down the positive path to a still lower node for examination of another specification, and so on until an endpoint or leaf was reached. At the endpoint of a tree, the recognition would be either assigned a main entry, A (author) or T (title), with designation of the item to occupy the entry, an author name or title of monograph, or be sent to the next selection rule for examination by that rule.

8. Finally, with respect to their application, the selection rules are arranged to be applied to a recognition in a given order: 1,2,3 ... Application starts at rule 1 and continues with the application of successive rules only until the specifications of entry are satisfied. As soon as an endpoint with an entry designation is reached, the selection process terminates. In the trees these endpoints are marked with the word STOP, indicating termination of the process.

A presentation of both the recognition rules and selection rules developed for each corpus of verbal main entry rules is given in the next section.

The Simulation Models

The characterization of the rules of entry selected from the AA and ALA cataloging codes was accomplished with the construction of two sets of formal rules for the two components comprising the entry selection process in each code, namely a recognition component and a selection component. In the case of the recognition component, the rules characterizing the component had no verbal counterpart in the cataloging codes. The cataloger's ability to recognize the presence of authorship is assumed, assisted with definitions of bibliographic terms and illustrations provided by the codes. The rules characterizing this component then were constructed out of those syntactic units identified as signalling devices for the presence of authorship. In the case of the selection component, the rules characterizing the component had their verbal counterpart in the AA and ALA cataloging codes, specifically, the corpus of entry rules selected from each code. Together, the two sets of formal rules, the recognition rules and the selection rules, formed a simulation model of the entry selection process in each code.

The two sets of formal rules developed for each code are given in Appendix B, pp. 328-71. It will be noted that the selection rules are given in both verbal and tree form. For the corpus of 19 AA rules of entry, there were seven rules in the recognition component and nine rules in the selection component. For the corpus of 36 comparable ALA rules of entry, there were seven rules in the recognition component and eleven rules in the selection component. With the exception of one rule, the recognition component for both sets of code rules was identical. The exception occurred with recognition rule 5 for the marker OF. In both models, the rule was constructed to recognize the marker after BY in the substring. Both rules functioned to recognize a physical and syntactic patterning of authorship associated with the use of two pairs of code rules: ALA 3B and AA 3A, and ALA 5C(4) and AA 19B.

However, in the case of the ALA recognition rule, the rule was context-free to permit the recognition of author names in substrings such as: By Lawrence T. Pereira With The Help of Joanne Taylor, and: By the Editors of Esquire, involving the use of the code's rules 3B and 5C(4), respectively. The AA recognition rule, on the other hand, was context-restricted, recognizing OF after BY in the context: BY ... WITH ... _______. This restriction functioned to recognize author names in substrings of the first type mentioned above, involving the use of the code's rule 3A, and avoided recognition of the marker in substrings of the second type, thus accommodating entry selection in accordance with the code's rule 19B.

In the case of the selection rules, there were nine similar, although not identical, rules in the two models. These included all the rules in the AA model and all but two in the ALA model. These rules shared the same construct: $\mathcal{Q}$ BY $\mathcal{X}Z$, etc., but differed in their branching, requirements, and specification of entry contained at the nodes and endpoints of the trees.

The two ALA selection rules that were not shared by the AA model were rules 6 and 7. These rules were constructed, in particular, to select entry for those compilations of essays, articles, etc. by various authors from a periodical or newspaper. In order to produce entry under the periodical or newspaper, as specified in the code's rule 5C(4), it was necessary to have both rules; the first, rule 6, to operate on recognitions with the structure: BY $\mathcal{X}$ of $\mathcal{Y}$, for example: By the Editors of Esquire, and the second, rule 7, to operate on recognitions with the structure: BY $\mathcal{X}$ AND $\mathcal{Y}$ OF $\mathcal{Z}$, for example: By The Cubs and Scouts of Boy's Life. At the same time, these rules were able to select entry for recognitions such as: By Lawrence T. Pereira With The Help Of Joanne Taylor, and: By William S. Bouton and Peter Vincent With The Assistance of Gene Miro, which require entry under the first author in accordance with the code's rule 3B, namely, by specifying for selection rule 6 that A (author) = $\mathcal{X}$, unless $\mathcal{Y}$ = a periodical (e.g. Esquire), then A = $\mathcal{Y}$, and for selection rule 7, that A (author) = $\mathcal{X}$, unless $\mathcal{Z}$ = a periodical (e.g. Boy's Life), then A = $\mathcal{Z}$.

It is, of course, obvious from the number of selection rules in both sets that several verbal rules were characterized with the construction of one selection rule. This was indicative of about all but four of the selection rules in each set, a point discussed in the following chapter (see pp. 207-11). In the latter cases, there was a one-to-one correspondence between a verbal rule and a selection rule. The reverse of this situation was indicative of one or two verbal rules in each corpus, where there were two or more selection rules to characterize a single verbal rule. One example is ALA 5C(4) above, which was characterized with selection rules 4,6 and 7. Additional points of comparison between the sets of experimental rules and the verbal rules they characterize will be discussed in Chapters VI and VII, which report a test of the adequacy of the simulation models to perform entry selection.

CHAPTER V

FIRST TEST OF ADEQUACY

This chapter is the first of two that will discuss the tests of adequacy performed on the two simulation models. Two tests of adequacy were conducted. The first test demonstrated the adequacy of the simulation models to generate correct main entry choices. This test, which was performed by the investigator, involved the application of each model to a random sample of Library of Congress proof slips. The procedures and findings of this test are presented and discussed in the present chapter. The second test demonstrated the adequacy of the models to accomplish selection of entry automatically. This test was performed in an experiment, which involved the application of the models to a specially selected problem set of English language monographs. The procedures and findings of this test will be presented and discussed in the next chapter.

Procedures

It was the main objective of the formal characterization to demonstrate that selection of entry can be accomplished in a completely objective, mechanical manner without the use of human intuition and foreknowledge. The strongest possible proof of the adequacy of the simulation models to accomplish this objective is to show that the models can be applied directly to the bibliographic data used in the selection of entry, and that on the basis of this application, appropriate main entry choices will be selected. For this purpose, two operational tests of adequacy were performed on each simulation model.

The necessity of conducting two separate tests of adequacy using the two different approaches was prompted out of practical considerations with respect to the nature of the materials developed in the study and the constraints of the investigation. One difficulty was that of preparing the large quantity of materials needed to conduct the desired testing. Altogether there were thirty-four formal rules developed in the study, sixteen of which characterized the corpus of AA rules of entry, eighteen of which characterized the corpus of ALA rules of entry. Practical application required that there be a complete set of rules, either AA or ALA, for each bibliographic item used in the testing. Since two groups of six persons (12 people) were expected to apply the rules to the same items, this meant that there must be 12 sets of rules for each item used for simultaneous testing. The bulk of the materials and cost of their duplication and collation, as well as the time required of the participants to apply them, indicated that testing on any scale involving more than fifty bibliographic items was totally unfeasible.

Finally, a difficulty related to this was that of acquiring the number and kind of bibliographic items needed to conduct the testing. A prerequisite was that the items be exemplary of the rules in each corpus. It was originally thought that the items could be acquired from the University of Illinois library stacks. However, on the basis of a preliminary investigation in which a sample of 100 to 200 Library of Congress proof slips was checked, it was found that this approach did not yield a sufficient number or variety of items.

By constructing separate operational tests of adequacy, it was possible to reduce the dimensions of these difficulties. For instance the models could be tested to determine their adequacy to produce appropriate main entry choices by applying them directly to the title and author statements recorded on Library of Congress proof slips. Use of the proof slips would eliminate the difficulty of acquiring bibliographic items and permit more rigorous testing of the models. With this approach, it would be possible to apply the models to larger random samples of bibliographic items representing the verbal rules characterized in the study. The test would also permit performance by the investigator. On the other hand, to test the adequacy of the models to accomplish selection of entry automatically, it would be necessary to apply the models directly to the title and author statements on the title pages of books. This test would also require performance of the participants mentioned. However, since size of the sample would not be as critical a factor in conducting the test, the difficulties in the way of acquiring the bibliographic items, preparing the materials, and administering their application would be greatly diminished.

In the following paragraphs, the first test of adequacy, which was performed by the investigator, is described. The procedures used to conduct the test are first discussed, followed by a separate section in which the findings resulting from these procedures are presented.

Selection of sample. — A random sample of Library of Congress proof slips was used in testing each model for its adequacy to produce appropriate main entry choices. Each sample was obtained from three weekly shipments received at the University of Illinois Graduate School of Library Science. The random sample used to test the simulation model representing selection of entry in the ALA cataloging code was from shipments 6, 9, and 11 received in the months of February and March 1966. The sample used to test the model representing selection of entry in the AA cataloging code was from shipments 25, 29, and 31 received in the period May - June 1969. Since the shipments were received under a blanket order placed with the Library of Congress, the contents were not biased by institutional or personal selection.

The decision to use these shipments for the samples was reached after an examination of the contents of the boxes. The necessity for this examination was two-fold. First, the models were constructed for application to the title pages of English-language monographs. This meant that only those proof slips whose entries were printed in English could qualify for use in the testing, and second, the rules comprising the models were constructed for application to monographs whose entries required to the use of the corpus of verbal rules characterized by the models. This meant that only those proof slips whose entries indicated the application of the verbal rules in question could qualify for use. Without prior examination of the contents of each box, there was no way of knowing whether the proof slips met these qualifications.

Examination of the shipments in question indicated that the entries were predominantly English-language and exemplary of the rules of entry in each corpus. Determination of rule use was first performed by the investigator. In the case of each entry, the code rule or rules were designated. This was followed by an independent examination of the entries that was performed by a cataloger with experience and knowledge of the rules in the two codes. Where an agreement was not reached as to rule use, the entries in question were examined by an additional independent source, and whenever possible, the titles were procured for examination.

The printing dates of the proof slips, indicated by the first two digits of the Library of Congress card number, were used as the indicator of the cataloging code employed to assign the main entry. In the case of the sample used to test the simulation model representing selection of entry in the ALA code, the proof slips, received in February and March 1966, had printing dates of (19)65 and (19)66. These proof slips were considered to be cataloged unquestionably under ALA rules of entry. In the sample used to test the model representing selection of entry in the AA code, the proof slips had printing dates of (19)68 and (19)69. Since the Library of Congress has adopted the policy of "superimposition" with respect to the application of the AA code, the latter dates were a partial assurance but, unfortunately, the only indicator of the use of that code.

The results of this examination are given in the following tables. Table 15 shows the distribution of the two samples in terms of the AA

Table 15

Distribution of Samples by AA and Comparable ALA Rules of Entry

AA Rule	LC Proof Slips	Percentage	ALA Rule	LC Proof Slips	Percentage
1A	1623	73.71	2	1505	69.29
			16	1	.05
			22C	15	.69
2A	20	.91	32	13	.60
2C	2	.09	32	13	.60
			32B	1	.05
3A	28	.14	3A	270	12.43
			3B	30	1.38
			4A(1)	25	1.[illegible]
3B1	330	10.45	4A(2)	7	.32
			4B(1)	1	.05
			6A	3	.14
			6B(1)	2	.09
3B2	10	.45	4A(2)	7	.32
4A	75	3.41	5A(1)	118	5.43
			5A(1c)	8	.37
			5B	9	.41
			5E(2)	4	.18
5A	104	4.72	5A(1)	118	5.43
			5A(1a)	2	.09
			6C	2	.09
5B	1	.05	5A(2)	1	.05
7A	8	.36	22B	7	.32
			23A	1	.05
			23B	1	.05
			24	2	.09
8A	4	.18	19D	27	1.24
8B	185	8.40	19E	109	5.02
9A	2	.09	6B(2)	3	.14
9B	1	.05	6B(2)	3	.14
14A	18	.82	20	56	2.58
			22A	32	1.47
14B	2	.09	20	56	2.58
15A	79	3.59	21A	97	4.47

Table 15 (continued)

AA Rule	LC Proof Slips	Percentage	ALA Rule	LC Proof Slips	Percentage
16A	3	.14	3E	2	.09
19B	15	.68	5C(4)	9	.41
			14A	3	.14
			25A	3	.14
			25C	1	.05
			26	4	.18
			28	2	.09

AA Rules of Entry	19
Sample of LC Proof Slips	2202
ALA Rules of Entry	36
Sample of LC Proof Slips	2172

and comparable ALA rules of entry. The two samples were comparable in size and, with some few exceptions, in rule distribution. The exceptions occurred chiefly with AA rules 3A, 3B1, 4A, 8A, and 14A and their ALA counterparts. However, in the case of two of these rules, AA 3A and 3B1, the differences are minimized when the entries for these rules are compared in combination with those of their counterparts. The distribution then results in 358 entries for the two AA rules as compared with 338 entries for their combined counterparts. Since the two rules are parts or sub-rules of the general AA rule 3, this comparison appears more meaningful.

With respect to the other rules, namely, 4A, 8A, 14A and their ALA equivalents, the question arises as to what significance the differences might have with regard to performance of entry selection by their respective models. While this question is pertinent, it will be discussed in connection with the findings, which are presented in the next section.

The distribution of the two samples by main entry type is shown in Table 16. The predominance of author main entry, in this case personal author, was predictable. It is of interest to note that the percentage of title main entries in both samples correlates closely with that determined in the two earlier samples presented in Chapter III (see Table 3, p. 37). In the earlier samples, it appeared that the differences in the distribution of this entry type were more attributable to the composition of the two samples than to the construct of the cataloging codes. In the case of AA 2A, for instance, 110 proof slips

Table 16

Distribution of AA and ALA Samples by Main Entry Type

	Author		Title		
	Proof Slips	Percentage	Proof Slips	Percentage	Total
AA	2121	95.87	81	3.67	2202
ALA	2126	97.88	46	2.11	2172

were assigned the entry in contrast with the 12 proof slips assigned that entry by its ALA counterpart. In the present samples, however, the distribution of entries for this rule, 3B2 and 4A, which together largely account for the use of this entry, is comparable with that of their ALA equivalents. The entries for all these rules were examined for their similarity with respect to type of material, wording of the title and author statements, etc.

The results of this examination indicated that in the present AA sample, the larger percentage of title main entries was attributable to the construct of the code and, in particular, to one rule, 4A. This is shown in Table 17, where identical or similar entries were paired by these set of rules for their main entry type. Calculation of x^2 (chi square) for the distribution in Table 16 provided a chi square value of 9.39. At one degree of freedom, x^2 indicates a value for p of .001, showing the distribution to be significantly different from what chance alone would produce, with more title main entries under AA and fewer under ALA than might be expected.

Application of the models. — Application of the simulation models rested on certain basic assumptions with respect to the arrangement of the bibliographic data recorded in the body of the entry of the proof slips. The assumptions similarly underscored application of the models to title page data, as stated in Chapter IV (see pp. 125-30, 157-61). They are stated below in the context of their specific relevance to the bibliographic data recorded on the proof slips in the two samples.

Table 17

Entries Paired by Sets of AA and Comparable ALA Rules of Entry for Main Entry Type

Rules	Proof Slips	Main Entry Type	
		Author	Title
AA 2A	8		8
ALA 32	8		8
AA 3B2	3		3
ALA 4A(2)	3		3
AA 4A	18	2	16
ALA 5A1	4	4	0
5A(1c)	7	6	1
5B	7	5	2
Totals			
AA Entries	29	2	27
ALA Entries	29	15	14

1. The bibliographic data on the proof slips which served as a basis for the application of the models consisted of the title and author statements, recorded in that sequence in the body of the proof slips. The recognition rules of both models identified these statements as two physically separate strings, the title statement as the initial or first string, and the author statement as the sub or second string, each string being composed of as many lines as there are in each statement. Application of the recognition rules rested on the abstraction of the entry as illustrated below:

Proof Slip

Gullixson, Thaddeus Frank, 1882 -
Down into the arena, and other sermons, by T. F. Gullixson.

Abstraction

#Down into the arena, and other sermons# } initial string
#by
T. G. Gullixson# } substring

2. To facilitate application of the recognition rules to author statements containing statements of translation, illustration, introduction, etc. in the body of the entry, it was assumed that these statements were displayed sequentially on the title page of the bibliographic item, as illustrated below:

Proof Slip

Harrison, Jeffery G.
A wealth of wildfowl, by Jeffery Harrison;
illustrated by Pamela Harrison.

Title Page Display

A wealth of wildfowl
by
Jeffery Harrison
illustrated by
Pamela Harrison

3. When an author statement was lacking in the body of an author entry, it was assumed that the statement was displayed on the title page of the item and was omitted from the body of the entry for reasons of its use or agreement with the form of the author name in the heading.[1] The recognition rules in accordance with this assumption identified the presence of the statement in the manner described above.

In cases where the title statement on the proof slip read: Poems, Novels, Collected Works, etc., it was assumed that the inflected form of the author's name was displayed on the title page of the bibliographic item. The recognition rules identified such statements as follows:

---

[1] See LC rule 3:6, U.S. Library of Congress, Descriptive Cataloging division, Rules for Descriptive Cataloging (Washington, D.C.: Government Printing Office, 1949), p. 14 and also its revision in U.S. Library of Congress, Processing Department, Cataloging Service, Bulletin no. 65 (August, 1964), p. 3 and AA rule 134A, The American Library Association and others, Anglo-American Cataloging Rules, North American Text (Chicago: American Library Association, 1967), pp. 96-97.

Proof Slip

De Noronha, Leslie.
Poems.

Abstraction

#Poems of Leslie De Noronha# } initial string

4. When an author statement was supplied in the body of the entry, recognized by its enclosure in brackets, it was an indication that the statement was absent from the title page of the bibliographic item.[2] The recognition rules accordingly did not identify the statement supplied in the entry. Similarly, when words were supplied in an author statement, the rules did not recognize these additions.

5. To facilitate uniform application of the recognition rules for the marker AND, certain assumptions were made with respect to the title page display of those author statements containing two or more author names, or one author name and the phrase "and others" in the body of the entry.

a. When an author statement contained two or three author names in the body of the entry, conjoined by "and" in brackets, it was assumed that the author names were displayed vertically on the title page without expression of conjoinment.

b. When an author statement contained three author names in the body of the entry, the first two being conjoined by a comma and the last one conjoined by "and," it was assumed that these names were displayed horizontally on the title page with the expressions of conjoinment noted.

---

[2] See LC rule 3:4, ibid, p. 10 and AA rule 132A1, ibid, p. 192.

c. When an author statement contained one author name followed by the phrase "and others" in brackets, indicating the presence of four or more author names on the title page of the item, it was assumed that there were four author names, and that these were listed vertically on the title page without any expressions of conjoinment.

Finally, in applying the models to the two samples, the investigator used the reverse side of each entry to record the recognition and to note the selection of entry. In recording the recognition, two lines were used for the two strings (initial and sub) as required by the rules. For example, for the entry:

Proof Slip

Crosetti, Frank.
Frank Crosetti's secrets of baserunning and
infield play. Illustrated by Bill Gallo.

application of AA recognition rules 1 and 2 (see Appendix B, pp. 329-30) resulted in the following recognition recorded on the reverse of the proof slip:

#Frank Crosetti's#
#Illustrated by Bill Gallo#

In applying the sets of selection rules to the recorded recognitions, the type of entry and item to occupy the entry were recorded in the lower right corner of the proof slip, followed by the number of the selection rule under which these were produced. For the recognition recorded above, application of the AA set of selection rules resulted in the following main entry selection:

A: Frank Crosetti's
3

Scoring. — The scoring procedure that was used in performing the test was to categorize each recognition as either "successful" or "unsuccessful," and each selection of entry as either "correct" or "incorrect." Although the distinction between the two sets of categories is undoubtedly subtle, it was believed that their use was justified. For instance, it was found that there were no "incorrect" or entirely wrong recognitions. While the rules may have failed to recognize an author name, or may have produced a "non" name (some word other than an author name), these were almost always found in association with other properly recognized names. On the other hand, application of the selection rules resulted in entries that were unquestionably "correct" or "incorrect" by the criterion for accuracy used in the test. It was for these reasons that the two sets of categories were established.

In the case of the recognition rules, a recognition was determined "successful" if application of the rules resulted in producing all and only all of the author names (writer, compiler, editor, artist etc.) recorded in the title and author statements on the proof slip. The production of a "non" name, or the failure to produce an author name in a recognition was determined an "unsuccessful" recognition. In an attempt to determine the nature and extent of these failures, the unsuccessful recognitions were analyzed by type. The production of "non" names was defined as "erroneous." Failure to recognize or produce author names was defined as "zero." The production of both "non" names and unrecognized author names was termed "erroneous and zero."

In the application of the selection rules to the recognitions, an entry choice was determined "correct" if it agreed or matched the entry choice with respect to both type of entry and heading on the Library of Congress proof slip. Lacking this agreement, an entry choice was determined "incorrect." The latter entries were similarly analyzed by type to determine the nature and extent of their occurrence. Incorrect entries were found to result from one of two errors: the choice of entry (author rather than title or vice versa), or the choice of author in the heading (an editor instead of the writer of the text, etc.).

In scoring the two samples, the investigator examined all the recognitions and main entry selections twice, once after the application of the models, and again after a lapse of six months. The correlation between the results was perfect. The number of successful and unsuccessful recognitions and the main entry selections which resulted were examined in various ways, as will be seen in the section on the findings.

Pretesting. --- Pretesting of the models and procedures was done in several stages with a number of individuals. After the models were initially developed, they were examined by a linguist of the faculty of the University of Illinois. Since the investigator had adapted the use of tree diagrams, symbols, etc. from the field of linguistics, it was believed that this examination was of critical importance to the study in general. As a result, certain changes were made with respect to the selection rules. In particular, it was found that both sets of these rules could be further reduced in number by collapsing one or two others contained in their sets. Certain minor changes were made in the abstract use of alphabetical symbols to represent author names in the selection rules. In place of a variety of these symbols, a core of five was chosen: $\mathcal{Q}$, $\mathcal{W}$, $\mathcal{X}$, $\mathcal{Y}$, Z. These were used as consistently as possible. For example, the first author name was always represented with $\mathcal{X}$; the $\mathcal{Q}$ and Z always represented elements that may or may not be present in a recognition, etc.

After making these changes, two sets of instructions were drafted to accompany the use of the two sets of rules in each model. These instructions and the models, along with a trial title page, were tried on two professional catalogers and two freshmen at the University of Illinois. As a result of this pretest, certain changes were made in the format of the recognition rules, and some additional explanations were inserted in the written instructions for both sets of rules.

After completing these changes, the procedure was repeated with a professional cataloger and a freshman at the University of Chicago. The results of this pretest indicated that the presentation of the models and the procedures for their use were satisfactory.

Findings

In this section, the findings resulting from the test of accuracy performed by the investigator are presented. The performance of the models is discussed under four numbered subheadings. In the first subheading, the findings related to the performance of the recognition rules in the models are presented. In the second subheading, those related to the performance of the selection rules are presented. These are followed by a third subheading in which the models are examined for their complexity. In a final subheading, the causal factors that were said to influence the performance of the models are examined.

1. Accuracy, as related to the performance of the recognition rules. — Since author names are the input or information on which the selection rules base the choice of entry, their successful identification is essential for the performance of that process. In the test conducted in the study, the number of proof slips, whose author names had been identified, indicated that the recognition rules in the ALA and AA models were highly successful, performing at the 93 and 95 per cent level respectively. The results of their application are presented in Table 18.

The comparability of their performance could be expected since six of the seven recognition rules in the two models were identical. The one pair of different or unique rules was constructed for the recognition marker OF, which was discussed in the preceding chapter (see p. 164). Consequently, the performance of the rules was essentially the performance of one set of recognition rules. This sharing of one set of rules indicated that the author recognition process is essentially identical for the application of the two cataloging codes.

Table 18

Recognitions Produced by Application of AA and ALA Recognition Rules to Library of Congress English Language Proof Slips

Rules	Proof Slips	Recognitions			
		Successful	Percentage	Unsuccessful	Percentage
AA	2202	2102	95.46	100	4.54
ALA	2172	2018	92.91	154	7.09

Table 19

Distribution of AA and ALA Samples by Recognition Rules*

Rules	Markers	Samples			
		AA	Percentage	ALA	Percentage
1	BY	2137	97.04	2144	98.71
2	'S,S'	14	.63	9	.41
3	OF/Life and	1	.04	3	.13
4	{OF FROM BETWEEN WITH}	15	.68	19	.87
5	OF	9	.41	19	.87
6	AND	465	21.11	319	14.68
7	WITH	23	1.04	31	1.42
Total		2664		2544	
Total Proof Slips Involved		2202		2172	

---

*With the exception of rule 5, the recognition rules in the AA and ALA models are identical.

The distribution of the samples with respect to the application of the rules is reported in Table 19. With one exception, the samples were comparable in rule use. The seeming discrepancy in the totals given for the proof slips results from the fact that most required the application of two or more recognition rules. As the footnote to the table indicates, with the exception of rule 5, the rules applied to the samples were identical. The frequency of rule use is of additional interest since it reveals something of the patterning of the bibliographic data on the proof slips in the samples. Proof slips with author statements having BY and AND, such as: "By Charles Kelley," "By Alice and William Goode," occurred most frequently. At the other extreme those with statements, such as "Life and Letters of Walt Whitman," were in a minority.

Type of recognition errors. –– Although the percentage of failure was statistically small, the nature and source of the recognition errors were considered of particular importance in determining the limitations as well as the potential of the rules.

Examination of the unsuccessful recognitions in the two samples indicated that over half were "erroneous," where application of the rules to the proof slips had resulted in the production of "non" author names, that is, words other than personal author names, such as: by eight leading assessors (the recognition marker and error in these and other illustrations will be indicated by underscoring). The results of this examination are given in Table 20. Of the remaining unsuccessful recognitions, all but three in the samples were "zero," where application of the rules resulted in the failure to recognize or

Table 20

Type of Unsuccessful AA and ALA Recognitions

Recognition Rules	Proof Slips	Recognitions					
		Erroneous	Percentage	Zero	Percentage	Erroneous and zero	Percentage
AA	100	55	55.00	45	45.00	0	0.00
ALA	154	100	64.93	51	33.12	3	1.95

produce an author name, as in: The Lenin reader. The three remaining recognitions, which were in the ALA sample, were "erroneous and zero," where the rules applied to produce both types of errors.

Source of recognition errors. — Although it is difficult to abstract the import of these particular types of errors, it was the opinion of the investigator that zero recognitions represented a more serious type of error, since it was found that the rules were largely limited in preventing their production in the samples. Table 21 shows the rules which applied to produce these recognitions. The discrepancy in the totals given for the proof slips in the samples results from the fact that in a few cases, two rules applied to produce two of these errors on a proof slip.

As the table shows, approximately 50 per cent of the zero recognitions in the two samples involved author statements to which the rules for the markers did not apply. These proof slips, 52 in the two samples, had statements where the author names were isolated or embedded in the title, as in: "The Lenin reader," or enclosed in brackets, a convention used in cataloging to indicate the absence of the name on the title page. In the case of four proof slips, author names occurred in association with the marker TO, as in: "[by] Ernest King, as told to Richard Viner," for which there was no recognition rule.

In the case of the embedded names, which accounted for 13 proof slips, it was found that author recognition would have to be accomplished at the expense of creating many ad hoc rules, which would be so

Table 21

Zero Recognitions Produced by Recognition Rules

Rules	Markers	Zero Recognitions			
		AA	Percentage	ALA	Percentage
1	BY	0	00.00	0	00.00
2	'S,S'	0	00.00	4	7.27
3	OF/Life and	0	00.00	0	00.00
4	OF FROM BETWEEN WITH	16	34.78	14	24.56
5	OF	0	00.00	0	00.00
6	AND	6	13.04	9	15.78
7	WITH	0	00.00	2	3.50
		24*	52.17	28*	49.12
Total		40		57	
Total Proof Slips Involved		45		54	

*Proof slips to which the recognition markers did not apply.

special and restrictive that they would jeopardize practical application. For instance, to recognize Lenin in the title above, there would have to be a rule, such as:

Recognize Z/__________reader

which would recognize any name in the environment of "reader." To accommodate recognition of Jack London in the title: "The Bodley Head Jack London," there would have to be a separate rule, which would:

Recognize Z / The Bodley Head __________

Similarly, each of the remaining titles required the construction of a special rule, since the position of the names and the contexts or environments in which they occurred were unique to the items.

The four proof slips, where the author names occurred in association with the marker TO, also required the construction of special rules to accommodate recognition. For instance, to recognize Richard Viner in "[by] Ernest King, as told to Richard Viner," there would have to be a rule which would recognize:

Z / as told to __________

Recognition of Walter Mapes in: "The Latin poems attributed to Walter Mapes [Collected and edited by] Thomas Wright," would require a rule which would:

Recognize Z / attributed to __________

and so on.

Where author names occurred in isolation on three of the proof slips, and had the dual function of an author and title statement, as in "Bernard Buffet," it was found that recognition was impossible unless the names could be checked against an authority name file. Even then, success in the use of this procedure would be restricted to works whose author names had been previously established.

Where author statements had been supplied by sources other than the title page, which involved 36 proof slips, recognition of the author names was impossible to accomplish, since the proof slips indicated that the names were not on the title pages of the items, but rather were obtained from the text or extraneous bibliographic sources.

The table shows that the remaining zero recognitions were chiefly produced by two recognition rules: AA and ALA 4 and 6, which were identical in the models. (For reference to these and other recognition rules discussed in the section, see Appendix B, pp. 329-35, 349-55). By adding further environmental restrictions in rule 4, the 30 zero recognitions produced by the rules could be prevented. These recognitions involved 30 proof slips, which would account for approximately 10 per cent of the unsuccessful recognitions produced in the two samples. This rule, which provided for recognizing author names in the contexts listed: "articles," "conversations," etc. failed when names were in contexts as, for example: "The speeches of Malcolm X," "The chronicles of J. Froissant," etc. With the inclusion of these contexts in the rule, successful recognition of the author names would be possible.

However, alteration of rule 6 would have the effect of providing for the recognition of words it was designed to avoid. For instance, by dropping the restriction placed on recognizing the marker AND in the environment of: WITH ... ___________, the 15 zero recognitions which occurred in author statements, such as:

By M. Almy
with E. Chittenden and P. Miller

would be avoided; however, the rule would then apply to produce erroneous recognitions in statements such as these:

By Emil Lathrop
with maps and drawings by E. J. Coleman

which are far more prevalent in title page displays. In the samples tested in the study, 100 to 150 proof slips would have been affected by dropping the restriction in the rule.

In contrast to the zero recognitions, it was found that all of the erroneous recognitions produced in the samples could be prevented, if the rules could "look up" a dictionary list or file of unacceptable words and phrases. For instance, if the restriction in rule 6 were dropped, as mentioned above, the erroneous recognition of and drawings (and similar phrases in the 100 to 150 proof slips which would be affected) could be prevented, if the rules had these words listed as "unacceptable" for recognition. For example, in addition to "edited", the rule would:

$$\text{Recognize AND} \neq \text{(but not) AND} \,/\, \underline{\qquad\qquad} \left\{\begin{array}{l}\text{edited}\\ \text{drawings}\\ \text{etc.}\end{array}\right\}$$

Similarly, in the illustration given earlier: "by eight leading assessors," a rule for the marker BY, which would:

$$\text{Recognize BY} \neq \text{BY}/\underline{\hspace{3cm}} \quad \left\{\begin{matrix}\text{eight}\\ \text{nine}\\ \text{ten}\\ \text{etc.}\end{matrix}\right\}$$

would prevent this erroneous recognition. In addition to numbers, descriptive phrases introduced with "some," "several," "important," "well known," etc. would also qualify for inclusion in the rule.

The rules involved in the production of these recognitions are given in Table 22. By far the largest number in the two samples resulted from the application of rule 6 for the marker AND. These occurred in connection with the rule's requirement that the marker be inserted at the end of all strings except the final string in the author statement. In order to accommodate the recognition of names in joint author statements, such as:

BY
L. K. Chan
Terri Kiossi

insertion of the marker was necessary. Consequently, when applied to statements such as:

BY
W. D. Park
second edition

the rule, in the absence of "looking up" or consulting a dictionary of "unacceptable" terms, produced the recognition: By W.D. Park [AND] second edition.

Table 22

Erroneous Recognitions Produced by Recognition Rules

Rules	Markers	Erroneous Recognitions			
		AA	Percentage	ALA	Percentage
1	BY	7	12.72	19	18.01
2	'S,S'	1	1.81	1	.90
3	OF/Life and	0	00.00	0	00.00
4	OF FROM BETWEEN WITH	0	00.00	1	.90
5	OF	2	3.63	14	12.61
6	AND	39	70.90	76	68.46
7	WITH	6	10.90	0	00.00
Total		55		111	
Total Proof Slips Involved		55		103	

The only unique rule in the two models was rule 5 for the marker OF. Although the table shows little difference in the performance of these rules, the erroneous recognitions they produced reflected their unique functions. To accommodate selection of entry in ALA code rules 3B and 5C(4), the ALA recognition rule functioned to recognize any occurrence of the marker in the substring, while selection of entry in comparable AA code rules required recognition of the marker only in the environment of WITH. Application to author statements, such as

BY

Lord Williams of Barnbugh

resulted in the ALA rule producing the erroneous recognition: of Barnbugh. The AA rule, however, was able to successfully block production of the error. The prevention of these errors gave some slight indication that the AA rule could be applied more successfully than the ALA rule. Whether this difference in the function of the rules affected successful application of the model's selection rules is discussed in section 2 of this chapter.

Thus, with the suggested alterations in the recognition rules, namely, the addition of environmental restrictions in AA and ALA rules 4 and the support of a dictionary list, the performance of the rules would be increased by three to five per cent in the AA and ALA samples respectively. This would indicate a performance level of 98 and 97 per cent in the respective samples. Only in the case of the 52 zero recognitions noted would author recognition be difficult, and in some cases impossible to achieve.

Summary. — In summary, the results of the application of the AA and ALA recognition rules to two random samples of Library of Congress proof slips indicated that the rules in the models were highly successful, performing at the 93 and 95 per cent level respectively. Comparability in the performance of the two sets of rules was expected since six of the seven rules in the models were identical.

Over half of the unsuccessful recognitions produced by the rules were "erroneous," where the rules produced non-author names. Of the remaining, all in the AA sample and all but three in the ALA sample were "zero," where the rules failed to recognize author names. The three remaining recognitions were "erroneous and zero," where application of the rules resulted in a combination of these errors.

Zero recognitions were considered a more serious type of error, since the rules were largely limited in preventing their production. However, all the erroneous recognitions could be prevented if there was provision for the rules to "look up" a dictionary of terms unacceptable for recognition.

Unsuccessful recognitions produced by the pair of different rules in the models reflected their unique functions. Although statistically, the differences in performance were fractional, application of the rules indicated that the AA recognition rule could be applied more successfully then the ALA rule.

With the suggested alterations in the recognition rules, it was found that their performance could be increased by three to five per cent, thus raising the performance level to 98 and 97 per cent in the AA and ALA samples respectively.

2. *Accuracy as related to the performance of the selection rules*.- It was proposed at the outset of the study that the experimental selection rules in the AA model would select correct main entry choices for a greater number of bibliographic materials than the experimental selection rules in the ALA model. This was because of certain factors related to the number and composition of their verbal counterparts in the two cataloging codes.

Application of the experimental selection rules in each model to the recognitions on the proof slips in their respective samples indicated that this prediction had no support. Table 23 shows the results of their application. Contrary to the prediction, both sets of rules were highly successful and performed equally well -- at the 97 per cent level.

The comparability in performance lent support to the conclusion that the two sets of experimental selection rules were equally powerful representations of their verbal counterparts. The finding also suggested that the rules were equivalent in their characterization of the requirements and objectives of the verbal selection process. In the following pages, the tenability of these conclusions is questioned through an examination of the models on a number of variables. The findings resulting from this examination are presented in Tables 24-37.

In contrast to the recognition rules, the selection rules in the models are similar but not identical. For instance, the majority of the rules in the models shared the same construct: $\mathcal{X} \overset{\&}{\text{AND}} \mathcal{Y}$ 'S,S',

Table 23

Main Entry Selections Produced by Application of AA and ALA Experimental Selection Rules to Recognitions

Rules	Proof Slips	Main Entry			
		Correct	Percentage	Incorrect	Percentage
AA	2202	2136	97.00	66	3.00
ALA	2172	2112	97.24	60	2.76

Table 24

Comparison of Author and Title Main Entries Produced by AA and ALA Experimental Selection Rules with Corresponding Rules of Entry

	Author				Title				
	Rules of Entry	Percentage	Experimental Rules	Percentage	Rules of Entry	Percentage	Experimental Rules	Percentage	Total
AA	2139	97.14	2125	96.50	63	2.86	77	3.50	2202
ALA	2149	98.94	2120	97.61	23	1.06	52	2.39	2172

Q BY XZ, etc ; however, the trees for the rules differed, in their branching, as can be seen from a visual examination, and in certain requirements and specifications of main entry contained at their nodes and endpoints. Another difference was the number of selection rules in each of the models, points which were discussed in the preceding chapter (see pp. 163-64). Further discussion of these points appears later in the section, however, it is mentioned here specifically since it serves to explain the difference in approach used to examine the performance of these rules and of the preceding recognition rules. Since six of the seven recognition rules in the models were identical, there was no basis for establishing the superiority of either model with respect to the characterization of the author recognition process underlying its verbal selection rules. However, differences in the selection rules noted above afforded a basis for evaluating their performance to determine if there was justification for choosing one model of entry selection over another.

Production of main entry type. --- The ability of the models to produce the required types of main entry (author and title) was seen as a factor of particular importance in determining the adequacy of their characterization of the verbal entry rules. The performance level of both models in the test pointed up their apparent success on this point. However, the distribution of the two samples by main entry type, which was shown earlier in Table 16, indicated that the incorrect entries produced by the models could be attributable to their failure to correctly assign the title main entry. Two to three per cent of the proof slips in the samples were title main entries, which was approximately the percentage of incorrect entries produced by the models. Failure on the part of both models to correctly assign this entry, for instance, would be indicative that the verbal requirements or conditions for that entry were inadequately characterized. A difference in the models on this point would be supportive evidence for establishing the superiority of one of the models.

To examine the models on this point, the correct and incorrect entries in the two samples were tabulated for their distribution by author and title. They were then compared with the printed entries produced by their verbal counterparts, reported in Table 16. The figures for the two sets of rules are combined in Table 24. As the results indicate, there were only slight differences in the models on this point. In general, both models had difficulty assigning correct title main entries. In the case of the AA model, a little over half or approximately 57 per cent of the printed title main entries in the sample were correctly assigned. In the case of the ALA model, about 41 per cent of those in its sample were correctly assigned. Although the AA model demonstrated more success in producing this entry type, support for the superiority of the model was not substantial.

Performance of rules. — The performance of the individual rules comprising the models was considered an additional factor in the way of determining the adequacy of a model's characterization of the verbal entry rules. In this connection, it was advanced that the adequacy of a model's characterization was associated with the number of selection rules that assigned correct main entries to the proof slips requiring their application. The assignment of correct entries was considered indicative of a rule's successful characterization of its verbal entry rules. The assignment of incorrect entries, on the other hand, was indicative of inadequacies in a rule's characterization of its counterparts. Thus, a model with 90 per cent of the rules assigning correct entries to the applicable proof slips would be considered more successful in characterizing its verbal entry rules than one with only 30 per cent of the rules assigning these entries. Although both models may have assigned the same number of correct and incorrect entries, the assignment of the former by a majority of a model's rules would indicate that it had successfully characterized the conditions of entry in more of its verbal rules.

To examine the models on this point, the proof slips in the two samples were separated first by rule number (the rule that assigned the main entry), and then by category (correct or incorrect). The results are reported in Tables 25 and 26, for the AA and ALA models respectively. As the tables show, there were again only slight differences in the models on this point. All the rules in both models assigned correct main entries. Of these, five rules assigned correct main entries to all their proof slips. These accounted for 55 per cent of the rules in the ALA model and 45 per cent of those in the AA model.

Table 25

Performance of AA Experimental Selection Rules as Applied to AA Recognitions

Rule	Proof Slips	Main Entry			
		Correct	Percentage	Incorrect	Percentage
1	4	4	100.00	0	00.00
2	10	7	70.00	3	30.00
3	4	4	100.00	0	00.00
4	1	1	100.00	0	00.00
5	15	15	100.00	0	00.00
6	19	7	36.84	12	63.16
7	4	4	100.00	0	00.00
8	2087	2055	98.47	32	1.53
9	58	39	67.24	19	32.76
Total	2202	2136	97.00	66	3.00

Table 26

Performance of ALA Experimental Selection Rules as Applied to ALA Recognitions

Rule	Proof Slips	Main Entry			
		Correct	Percentage	Incorrect	Percentage
1	3	3	100.00	0	00.00
2	4	4	100.00	0	00.00
3	2	2	100.00	0	00.00
4	3	3	100.00	0	00.00
5	19	17	89.47	2	10.53
6	13	12	92.31	1	7.69
7	6	4	66.66	2	33.33
8	18	2	11.11	16	88.89
9	2	2	100.00	0	00.00
10	2068	2046	98.94	22	1.06
11	34	17	50.00	17	50.00
Total	2172	2112	97.24	60	2.76

The remaining rules in the two models assigned incorrect as well as correct main entries.

It is of interest to note that the pairs of similarly constructed rules in the models were comparable in both performance and frequency of use. The results of their application, reported separately in the tables, are combined in Table 27. Four of the five rules that assigned correct entries to all their proof slips were in pairs. Of those assigning incorrect entries, three were in pairs.

If it can be assumed that the adequacy of a model's characterization is associated with the number of rules assigning correct entries to the proof slips, then the evidence produced in this examination would indicate that the two models were almost equally successful in characterizing their verbal entry rules. The performance of the pairs of similarly constructed rules in the models would further indicate that they had both succeeded and failed to characterize a number of verbal rules considered comparable in the study.

Characterization of verbal rules. --- The number of verbal entry rules successfully characterized by a model was considered a major factor in determining the adequacy of its characterization. If the models were in fact almost equally successful in characterizing their verbal entry rules, as the evidence above indicated, then it could be expected that they had succeeded in characterizing a comparable number of these rules. To test the models on this point, the proof slips assigned main entries by the model's selection rules were examined for

Table 27

Performance of Similar AA and ALA Experimental Selection Rules

Rules		Proof Slips		Main Entry							
				Correct				Incorrect			
AA	ALA	AA	ALA	AA	Percentage	ALA	Percentage	AA	Percentage	ALA	Percentage
1	1	4	3	4	100.00	3	100.00	0	00.00	0	00.00
2	2	10	4	7	70.00	4	100.00	3	30.00	0	00.00
3	3	4	2	4	100.00	2	100.00	0	00.00	0	00.00
4	4	1	3	1	100.00	3	100.00	0	00.00	0	00.00
5	5	15	19	15	100.00	17	89.47	0	00.00	2	10.53
6	8	19	18	7	36.84	2	11.11	12	63.16	16	88.89
7	9	4	2	4	100.00	2	100.00	0	00.00	0	00.00
8	10	2087	2068	2055	98.47	2046	98.94	32	1.53	22	1.06
9	11	58	34	39	67.24	17	50.00	19	32.76	17	50.00
Total		2202	2153	2136		2096		66		57	

their application of verbal main entry rules. The results of this examination are presented in Tables 28 and 29, for the AA and ALA models respectively. By following the columns marked "correct" under the experimental rules, it is possible to determine which as well as how many verbal rules were successfully characterized. Using the same procedure for the columns marked "incorrect," the number of verbal rules that were inadequately or unsuccessfully characterized can also be determined.

As an examination of the columns shows, only eight or 42 per cent of the verbal entry rules in the AA corpus were successfully characterized, as compared to twenty-two or 61 per cent of the comparable verbal rules in the ALA corpus. With the exception of two ALA rules, 4B(1) and 5A(1a), the evidence indicates that the remaining verbal rules in each corpus were partially characterized; this is, the proof slips representing the application of these verbal rules were assigned both correct and incorrect entries by the experimental rules. The two ALA rules, 4B(1) and 5A(1a), were the only verbal rules that were unsuccessfully characterized. In these cases, all the proof slips representing their application were assigned incorrect main entries.

By reading down the tables, it is of interest to note that most of the experimental rules in the models assigned entries to proof slips representing the application of several verbal rules. For instance, the proof slips assigned entries by experimental rule AA 8 represented the application of the 19 verbal rules in the AA corpus. As was pointed out in the preceding chapter (see pp. 164-65), the

Table 28

Correspondence of Main Entries* Produced by AA Experimental Selection Rules with AA Rules of Entry

AA Rules of Entry	AA Experimental Selection Rules: Main Entries																		Total	Percentage
	1		2		3		4		5		6		7		8		9			
	C	I	C	I	C	I	C	I	C	I	C	I	C	I	C	I	C	I		
1A			7		4				15						1569	12		16	1623	73.71
2A				2												2	16		20	.91
2C																1	1		2	.091
3A												4			24				28	1.26
3B1												1			328			1	330	10.45
3B2											7					2	1		10	.45
4A											2	7	4		50	7	5		75	3.41
5A											1				95	2	6		104	4.72
5B															1				1	.045
7A															4	2		2	8	.36
8A															2		2		4	.18
8B					1						1				183				185	8.40
9A							1								1				2	.091
9B															1				1	.045
14A	5														11	2			18	.72
14B															2				2	.091
15A															79				79	3.59
16A															3				3	.14
19B				1											5	1	9		15	.68

*Designated as "C" for correct and "I" for incorrect.

Table 29

Correspondence of Main Entries* Produced by ALA Experimental Selection Rules with ALA Rules of Entry

ALA Rules	Experimental Selection Rules: Main Entries																						Total	Percentage
	1		2		3		4		5		6		7		8		9		10		11			
	C	I	C	I	C	I	C	I	C	I	C	I	C	I	C	I	C	I	C	I	C	I		
2			2		1				9	1	3	1							1470	7		11	1505	69.75
3A													2			6			262				270	12.43
3B									1		5		2						22				30	1.38
3E																			2				2	.09
4A(1)																10	2		13				25	1.16
4A(2)															2				4		1		7	.32
4B(1)																						1	1	.05
5A(1)																			118				118	5.43
5A(1a)																				2			2	.09
5A(1c)																			7		1		8	.37
5A(2)											1												1	.05
5B																			7	1	1		9	.41
5C(4)			1						1		2			2					1	2			9	.41
5E(2)	1																		3				4	.18
6A									2											1			3	.14
6B(1)					1				1														2	.09
6B(2)							3																3	.14
6C																			2				2	.09
14A																			3				3	.14
16																			1				1	.05
19D																			24	1		2	27	1.25
19E					1						1								107				109	5.02
20	2									1									51	2		1	57	2.54
21A																			97				97	4.47
22A																			32				32	1.48
22B																			6	1			7	.32
22C									3										6	4		2	15	.69

Table 29 (continued)

ALA Rules	Experimental Selection Rules: Main Entries																						Total	Percentage
	1		2		3		4		5		6		7		8		9		10		11			
	C	I	C	I	C	I	C	I	C	I	C	I	C	I	C	I	C	I	C	I	C	I		
23A																			1				1	.05
23B																			1				1	.05
24																			2				2	.09
25A																			3				3	.14
25C			1																				1	.05
26																			4				4	.18
28																			2				2	.09
32																				1	12		13	.60
32B																					1		1	.05

Total Main Entries Involved 2172

---

*Designated as "C" for correct and "I" for incorrect.

experimental rules with few exceptions characterized patterns of authorship associated with two or more verbal rules. In some cases, the tables indicate the reverse, where the proof slips representing the application of a verbal rule are distributed over several experimental rules. For example, reading across Table 28, it will be seen that the 1623 proof slips representing the application of AA 1A were assigned entries by five experimental rules. In such cases, the patterns of authorship associated with the one verbal rule were characterized by several experimental rules.

Although the evidence produced on this point indicated that the ALA model was considerably more successful in characterizing its verbal counterparts, other considerations of the data showed the models to be almost equally successful on this point. For instance, when the 36 verbal entry rules in the ALA corpus were correlated with the 19 comparable verbal rules in the AA corpus, the evidence indicated that the ALA model had successfully characterized only seven or 37 per cent of the comparable rules of entry. This compared with the eight or 42 per cent that were successfully characterized by the AA model. Table 30 shows the models' assignment of main entries correlated by their application of comparable verbal rules. It will be seen that AA 1A, for example, was partially characterized, as evidenced by the assignment of correct and incorrect main entries to the proof slips representing its application. The comparable verbal ALA rules, which comprised rules 2, 16, and 22C, similarly showed partial characterization. While the proof slip representing the application of rule 16, for instance, was

Table 30

Main Entries Correlated by AA Rules of Entry and Comparable ALA Rules of Entry

AA Rule	LC Proof Slips	Correct	Incorrect	ALA Rule	LC Proof Slips	Correct	Incorrect
1A	1623	1594	29	2	1505	1486	19
				16	1	1	
				22C	15	9	6
2A	20	16	4	32	13	12	1
2C	2	1	1	32	13	12	1
				32B	1	1	
3A	28	24	4	3A	270	264	6
				3B	30	29	1
				4A(1)	25	15	10
3B1	330	325	5	4A(2)	7	7	
				4B(1)	1		1
				6A	3	2	1
				6B1	2	2	
3B2	10	8	2	4A(2)	7	7	
4A	75	67	8	5A(1)	118	118	
				5A(1c)	8	8	
				5B	9	8	1
				5E(2)	4	4	
5A	104	99	5	5A(1)	118	118	
				5A(1a)	2		2
				6C	2	2	
5B*	1	1		5A(2)	1	1	
7A	8	5	3	22B	7	6	1
				23A	1	1	
				23B	1	1	
				24	2	2	
8A	4	4		19D	27	24	3
8B*	185	185		19E	109	109	
9A*	2	2		6B(2)	3	3	
9B*	1	1		6B(2)	3	3	
14A	18	15	3	20	57	53	4
				22A	32	32	
14B	2	2		20	57	53	4
15A*	79	79		21A	97	97	
16A*	3	3		3E	2	2	

Table 30 (continued)

AA Rule	LC Proof Slips	Correct	Incorrect	ALA Rule	LC Proof Slips	Correct	Incorrect
19B	15	13	2	5C(4)	9	5	4
				14A	3	3	
				25A	3	3	
				25C	1	1	
				26	4	4	
				28	2	2	
Total 19	2202	2136	66	Total 36	2172	2112	60

*Sets of verbal entry rules successfully characterized by the models.

assigned a correct main entry, those representing the application of 2 and 22C, which also comprised the comparable rule, were assigned both entries.

This presentation of the data also revealed another point of comparison in the performance of the models. As the footnote to the table explains, the asterisks indicate those pairs of comparable verbal rules that were successfully characterized by the models. Of the eight rules in the AA corpus and the seven in the comparable ALA corpus that were successfully characterized by the models, six were in pairs. With three exceptions, the remaining pairs of comparable verbal rules were all partially characterized by the models.

In short, when the verbal rules of entry in each corpus were correlated according to their comparability, the evidence indicated that an almost equal number were successfully characterized by the two models. Similarly, an almost equal number were partially characterized by the models. Moreover, the data showed that most of these were found to be paired sets.

Source of entry errors. —— In connection with these findings, it was important to the study in general and to the selection of a model in particular, to determine how each of the models failed in the characterization of their verbal counterparts. If incorrect entries resulted from incomplete characterization of the verbal entry rules, namely, by failing to examine or provide for patterns of authorship associated with the rules, or were rather attributable to verbal conditions of entry too obscure for characterization, then it was advanced that the model whose selection rules were merely incomplete would be a better representation of its verbal counterpart and consequently superior to one whose rules were deficient. While "incomplete" rules would, in general, permit revision, "deficient" rules would be impossible or, at best, difficult to rectify. An incorrect entry resulting from the failure on the part of a rule to examine an author name in an environment such as "dramatized by," for instance, could be rectified by incorporating the environment in the rule. An incorrect entry resulting from the failure of a rule to distinguish a work "issued without text or a text of obviously secondary nature,"[3] on the other hand, would be impossible to rectify. Limitations of the study as well as those of characterization would prevent rectification of the rule's deficiency.

To determine if the incorrect entries assigned by the models were attributable to incomplete characterization of their verbal entry rules or, rather, to conditions of entry in the verbal rules that were too obscure for characterization, the 60 proof slips in the ALA sample and the 66 in the AA sample were examined for their entry error. The verbal entry rules applicable to the proof slips were then examined for their characterization in the model's selection rules. The results of this examination are summarized in the following paragraphs under three letter subheadings.

---

[3] See ALA rule 19A in: The American Library Association, Division of Cataloging and Classification, ALA Cataloging Rules for Author and Title Entries, 2d. ed., ed. by Clara Beetle (Chicago: American Library Association, 1949), p. 40.

A. Incorrect entries due to unsuccessful recognitions. —— Examination of the incorrect entries assigned by the models showed that over half, 61 per cent in the AA sample and 70 per cent in the ALA sample, were due to unsuccessful recognitions, which resulted from application of the recognition rules. The number of proof slips and the type of unsuccessful recognition, which accounted for the entry error are reported in Table 31. In both samples, the majority of these recognitions were "zero," where the author names had been unrecognized, due to either their isolation on the proof slips, such as "Bernard Buffet," their embedding in the title as, for example: "Lenin" in "The Lenin reader," their association in environments such as: "conversations of," "speeches of," etc., or their addition to the title statement, indicated by brackets. Without the recognition of the names in these statements, the selection rules in the models were unable to specify the main entry which involved their assignment.

The few "erroneous" recognitions, which resulted in the assignment of incorrect entries, were due chiefly to the insertion of the recognition marker AND at the end of the strings in the substring in order to provide for the recognition of names in statements of joint or shared authorship. Thus, with statements such as:

BY
J. L. Gayler
Irene Richards
J. A. Morris
2d. ed.

Table 31

Number of Incorrect AA and ALA Main Entry Selections
Produced by Unsuccessful Recognitions

Recognitions	Incorrect Main Entry Selections AA	Incorrect Main Entry Selections ALA
Erroneous	6	8
Zero	34	33
Erroneous and Zero	0	1
Total	40	42
Total		
Unsuccessful Recognitions	100	154
Incorrect Main Entry Selections	66	60

where the recognition rules indicated a pattern of four authors: By J. L. Gayler [AND] Irene Richards [AND] J. A. Morris [AND] 2d. ed., the selection rules in the models specified an incorrect main entry type (title instead of author). By making alterations in the recognition rules suggested earlier (see pp. 193-94), all of the erroneous recognitions and approximately half of the zero recognitions reported in the table could be assigned correct main entries by the models.

At the same time, the table shows that most of the unsuccessful recognitions produced by the recognition rules were assigned correct main entries by the selection rules. Examination of these recognitions, the remaining 60 in the AA sample and 112 in the ALA sample, indicated that there were two explanations for this phenomenon: the selection rules either blocked, or did not require the use of the recognition error for assignment of the main entry.

The rules blocked the use of recognition errors, namely, by examining the syntactic patterning of the data in the recognitions. For instance, in the 26 proof slips in the two samples with erroneous recognitions involving the marker BY (see Table 22, p. 195) where the author statements had the pattern: "by a Chinese artist" or "by the editors," the selection rules in the models blocked assignment of the author main entry by examining the statements for the presence of the definite and indefinite articles (a, the), and specifying that the main entry could not be A (author) when the statement was: "by a ..." or "by the ..."

The selection rules, on the other hand, did not require the use of certain recognition errors due to the physical patterning of the data in the recognition. For instance, in the case of the four proof slips in the samples with zero recognitions involving the marker TO (for which there was no recognition rule), such as: [by] Ernest King as told to Richard Viner (the marker and the recognition failure underlined), the selection rules in the models specified an author main entry under "King." In accordance with the verbal requirements of entry in AA 16A and its comparable ALA rule 3E, the main entry in these cases is under the first author name. Consequently, the recognition of the second author name was not necessary for the selection rules to assign correct main entries to these works.

Similarly, in the case of the six erroneous recognitions in the AA sample involving the marker WITH, where non-author names occurred in such statements as: "by L. Dudley Stamp with 47 colour photos. ...," the selection rules specified an author main entry under "Stamp." The pattern of the statement: BY ... WITH ..., in these cases, incorrectly indicated the works of shared authorship, However, the verbal rule, AA 3A, for such works with this pattern dictated entry under the first author name unless the name in secondary position is given typographical prominence. Consequently, despite the non-author name in the pattern, the experimental rules were able to assign correct main entries.

B. Incorrect entries due to incomplete characterization. — Examination of the remaining incorrect entries by the models indicated that 14 in the AA sample and 9 in the ALA sample, or approximately 21 and 15 per cent in both samples, were due to incomplete characterization of the verbal rules of entry in each corpus. In Table 32, the distribution of these entries is presented in terms of the experimental rules that assigned the entries and the corresponding verbal entry rules that were applicable to them. In the case of each proof slip, the type of entry error, either in the main entry type (author or title) or the heading (author name), is noted.

As can be seen, there were four experimental rules in each model that were involved in the assignment of these entries. Three of these were pairs of similar rules, namely: (AA 6, ALA 8), (AA 8, ALA 10), and (AA 9, ALA 11). In turn, the incorrect entries assigned by these rules represented the application of six verbal entry rules in each corpus. Of these verbal rules, the correlation in Table 30 shows that half were considered comparable sets of rules in the study, namely: (AA 1, ALA 2), (AA 3A, ALA 3A, 4A(1)), and (AA 7A, ALA 22B).

In the case of five rules, three AA and two ALA, incomplete characterization resulted from a failure to provide for the recognition of certain environments associated with their patterns of authorship. For example, in the three AA rules, 1A, 7A, and 19B, the proof slips representing their application had the patterns:

(1) preface by $\mathcal{X}$ written by $\mathcal{Y}$

(2) by $\mathcal{X}$ $\left\{\begin{matrix}\text{adapted}\\ \text{simplified}\end{matrix}\right\}$ by $\mathcal{Y}$

(3) concordance to $\mathcal{X}$'s poems

Table 32

Number and Type of Main Entry Errors Produced by Incomplete Characterization

Experimental Rules	Proof Slips	Entry Error		Rules of Entry
		Type (Author or Title)	Heading	
AA:				
2	1	1		19B
6	1	1		3A
	3	3		4A
8	6		6	1A
	2		2	7A
9	1	1		3B1
Total 4	14	6	8	6
ALA:				
5	1		1	20
8	3	3		3A
	2	2		4A(1)
10	1		1	22B
11	1	1		2
	1	1		19D
Total 4	9	7	2	6

In (1) and (2), failure to provide for the recognition of the environments "written," "adapted," and "simplified" in experimental rule 8 resulted in the rule's assignment of the author entry to $\mathcal{X}$.[4] In the absence of providing for the recognition of the environment "concordance to" in (3), experimental rule 2 incorrectly assigned an author main entry.

In the case of the two ALA rules, 2 and 22B, the proof slips representing their application had similar patterns and environments:

(4) written and illustrated by $\mathcal{X}$

(5) text by $\mathcal{X}$ adapted from ... by $\mathcal{Y}$

In (4) only the environment "illustrated" was recognized in the model's rule, thus causing rule 11 to incorrectly assign a title main entry. In (5), failure to provide for $\mathcal{X}$ and $\mathcal{Y}$ occurring in their environments resulted in rule 10 assigning the author entry to $\mathcal{X}$.

In the case of the remaining rules noted in the table, the environments associated with their patterns of authorship were incorporated but not sufficiently examined in the trees of the experimental rules to permit the assignment of correct main entries. For example, in the three AA rules, 3A, 3B1 and 4A, which involved works of multiple authorship, the proof slips representing their application had such patterns as:

---

[4] To provide for the recognition of these environments, the experimental rules, in general, would be of the form: A is $\mathcal{X}$ unless $\mathcal{Y}$ / adapted by ___________, simplified by ___________, written by _________, then A is $\mathcal{Y}$. Similarly, in the case of the experimental rule 2, A is $\mathcal{X}$, unless $\mathcal{X}$ / concordance to ___________, then T. In the event of: concordance to $\mathcal{X}$'s poems by $\mathcal{Y}$, the rule would reject the recognition, and an author main entry would be assigned to $\mathcal{Y}$ with application of experimental rule 8 in the AA model and 10 in the ALA model.

(6) By $\mathcal{X}$ and associates: $\mathcal{Y}_1$ and $\mathcal{Y}_2$ and $\mathcal{Y}_3$

(7) By $\mathcal{X}$ editor and contributors: $\mathcal{Y}_1 \overbrace{\text{and } \mathcal{Y}_2 \text{ and } \mathcal{Y}_3 \text{ and } \mathcal{Y}_4}^{\text{[and others]}}$

where failure to examine $\mathcal{Y}_1$ in (6) and $\mathcal{X}$ and $\mathcal{Y}_1$ in (7) resulted in experimental rules 6 and 9 assigning title main entries.[5] In the ALA rules, 3A and 4A(1), the proof slips representing their application had patterns and environments similar and, in some cases, identical to those given for the AA rules above. Failure to examine these environments in turn caused the model's similar experimental rule to assign incorrect main entry choices.

By making the adjustments noted, either in the way of incorporating or examining these environments in the experimental rules, all 23 proof slips reported in the table could be assigned correct main entries by the models.

---

[5] To provide for an examination of these environments, the nodes in the trees of the rules would contain the specifications: Is $\mathcal{X}$ in the environment of "editor"? (yes, no) Is $\mathcal{Y}_1$ in the environment of "associates"? and so on until all likely environments were exhausted and the entry specification was reached.

C Incorrect entries due to deficient characterization — The remaining proof slips, 12 in the AA sample and 9 in the ALA sample, were assigned incorrect entries due to conditions of entry in the verbal rules that were either conflicting or too obscure for formal characterization. In Table 33, the distribution of these proof slips is presented in terms of the experimental rules that assigned the entries and the corresponding verbal entry rules that were applicable to them. For each proof slip, the type of entry error is noted.

In the case of both models, an almost equal number of experimental rules was applicable to the proof slips. All the experimental rules and, with few exceptions, the verbal entry rules are also found in Table 32. This evidence showed a number of the experimental rules in the models to be both deficient and incomplete in the characterization of these verbal rules. Since the verbal rules had two or more conditions for assigning the main entry, this finding was not unexpected. For instance, if a rule had two verbal conditions of entry, one could be found to merely lack characterization, while the other could be too obscure for characterization.

Examination of the 21 proof slips noted in the table showed the patterns of the author statements to be insufficient for the purposes of assigning the main entry type specified by the verbal rules. All the proof slips in the ALA sample, and all but two in the AA sample involved the application of rules for works of multiple authorship, which were considered comparable in the study. The conditions of

Table 33

Number and Type of Main Entry Errors Produced
by Deficient Characterization

Experimental Rules	Proof Slips	Entry Error		Rules of Entry
		Type (Author or Title)	Heading	
AA:				
2	2	2		2A
6	3	3		3A
8	5	5		4A
	2	2		5A
Total 3	12	12		4
ALA:				
8	5	5		4A(1)
	3	3		3A
10	1	1		5B
Total 2	9	9		3

entry in these rules, which prevented characterization, are discussed briefly below.

Five, or almost half, of the proof slips in the AA sample involved the application of AA 4A, where the pattern of the author statement was:

(8) edited by $\mathcal{X}$

(9) editor: $\mathcal{X}$, or by $\mathcal{X}$ editor

Although AA 4A specifies an author main entry when the editor is named on the title page, entry under title prevails, when a publisher is named in the title of the work, or the editor is not primarily responsible for the existence of the work. In the case of the entries affected, such as:

(10) The International encyclopedia of science.
Edited by James R. Newman.

(11) Attorney's handbook on charitable giving.
William C. Golden, editor.

editorial responsibility for the existence of the work was lacking, and entry was under title. From the viewpoint of characterization, however, this condition was too obscure to be formalized, thus resulting in the model's rule 8 incorrectly assigning an author main entry. Possibly, with resource to catalog files, which could supply a "history" of entries for the edition in process, or textual data (preliminary pages, prefatory remarks, etc.), the condition of responsibility in this rule might be partially characterized. However, the likelihood of ever achieving uniform application of 4A is underscored in the rule's footnote 6a.[6] In the absence of relevant information in the text, the cataloger is instructed to judge the nature and extent of responsibility exercised by the editor. Positive answers, for instance, with respect to the editor's conception, plan and development of the work would indicate responsibility warranting main entry under the editor. A negative or doubtful response, on the other hand, would require assignment of a title main entry. Consequently, textual data, which was beyond the scope of the study, would provide only limited characterization of the rule.

The one proof slip in the ALA sample which involved ALA 5B similarly had the pattern:

(12) Edited by x

Although ALA 5B specifies a title main entry for encyclopedias and dictionaries, entry under author prevails when these publications are better known by the names of their editors. In the case of the entry affected:

(13) The Concise encyclopedia of modern
world literature.
Edited by Geoffrey Grigson.

---

[6] U.S. Library of Congress, Processing Department, Cataloging Service, Bulletin No. 88 (January, 1970), 1.

the rule applied to assign a title main entry. The demand of the rule in the way of identifying type of publication and the association of a name with a publication indicated that characterization was difficult, and at best only partially achieved. With the recognition of the words identifying the publications, i.e. "encyclopedia" and "dictionary," successful entry of works such as the above might be accomplished.[7] With resource to catalog files, further characterization might also be achieved.

Three of the 12 proof slips in the AA sample involved the application of AA 3A, where the pattern of the author statements:

$$\text{(13) By } \mathcal{X}\overbrace{\text{and } \mathcal{Y}_1 \text{ and } \mathcal{Y}_2 \text{ and } \mathcal{Y}_3}^{\text{[and others]}}$$

indicated the involvement of four or more authors in the production of the works. For works of shared authorship, AA 3A specifies an author main entry when principal responsibility can be determined, e.g. by wording and typography. This condition is operative regardless of the number of authors named in the author statement. Consequently, in the entries affected, such as:

---

[7] For instance, in recognizing "encyclopedia" and "dictionary" in the initial string, the model's selection rule 10 would act on the recognition: Is BY $\mathcal{X}$ in the environment: "encyclopedia or dictionary ... edited?, "and in confirming the environment, specify a title main entry. Difficulties, however, are encountered in such cases as: The Complete Encyclopedia of Ice Hockey . . edited by Zander Hollander & Hal Bock, where entry is under editor presumably because the publication is the work of the individuals and lacks the characteristic of continued publication.

Judo,

by Sadaki Nakabayashi [and others]

the rule applied to assign an author main entry.

However, when this condition is determined absent, and there are four or more authors in the author statement, AA 3B2 assigns a title main entry. The applicability of the same pattern of authorship to two rules specifying conflicting entry types resulted in the model's rule 6 incorrectly assigning title main entries. Characterizing the condition of responsibility in this pattern was partially achieved by providing for an examination of the wording and typography of the author statement. However, the demands of the rules in the way of using catalog files, examining textual data, etc. indicated that the condition was too obscure to permit complete characterization in the study.

In the case of the eight proof slips in the ALA sample involving the application of ALA 3A and 4A(1), the patterns of the author statements were identical to those discussed in connection with AA 3A above. However, assignment of the main entry was contingent upon the type of multiple authorship. If the works were those of joint authorship, ALA 3A would apply to assign an author main entry to the person determined chiefly responsible. However, if chief responsibility could not be determined, and there were four or more authors in the statement, ALA 4A(2) would apply to assign a title main entry.

In five of the proof slips affected, such as:

> Electron microscopy crystals,
> By P. B. Hirsch [and others]

examination of the items and the "notes" on the proof slips indicated that authorship was composite (individual parts by selected authors), with principal responsibility attributed to the first named author. Consequently, ALA 4A(1) applied to assign the main entry. In the case of the other proof slips, examination of the items indicated joint authorship, in which case ALA 3A applied to assign the main entry to the first named author.

As in the case of AA rules, 3A and 3B2, the applicability of the same general pattern of authorship to three rules specifying conflicting entry types resulted in the ALA experimental rule 8 incorrectly assigning title main entries. However, while characterization of the AA rules was partially achieved by providing for an examination of the wording and typography of the author statement, it was impossible, not just difficult, to achieve characterization of the types of multiple authorship differentiated by the ALA rules. Failing this characterization, the experimental rule 8 regarded this pattern as evidence of composite authorship with chief responsibility indeterminable.

Of the remaining four proof slips, which were in the AA sample, two involved the application of AA 2A, and two required application of AA 5A. In those involving AA 2A, the pattern of the author statement was:

(14) $\mathcal{X}$'S . . .

where the $\mathcal{X}$ indicates the presence of an author name. The applicability of AA 2A was determined upon examination of the items. Bibliographic sources as well as the texts indicated that authorship was unknown or, at best, uncertain. Consequently, the rule applied to assign title main entries. However, the presence of an apparent author on the title pages, as in:

Jawaharlal's discovery of America

resulted in the model's assignment of author main entries. Although there was no proof slip in the ALA sample with an identical pattern involving the application of the comparable rule, ALA 32, incorrect assignment of the author main entry would similarly have resulted. As in the case of the previous rules involving the characterization of principal responsibility, characterizing the condition of unknown authorship was only partially achieved, specifically, by examining the author statement for the presence of the articles (a, the), or by noting the absence of the author statement. The demand of the rules to consult and interpret information extrinsic and/or intrinsic to the works indicated that the condition was too obscure to completely characterize in the study.

The remaining two proof slips, involving the application of AA 5A, had author statements with the pattern:

(15) By $\mathcal{X}$ BY $\mathcal{Y}$

Since the proof slip data indicated that the item:

> Two stories: Argo, or The voyage of a balloon, by Andreas Embiricos. We Killed Mangy-Dog, by Luis Bernardo Honwana

was a compilation of independent works by different authors, with a collective title, and no compiler or editor, the rule applied to assign a title main entry. However, when the same compilation lacks a collective title, AA 5B applies to assign an author main entry. Here, as in the case of AA 3A and 3B2, discussed earlier, the applicability of the same pattern of authorship to two rules specifying conflicting entry types resulted in the model's rule 8 incorrectly assigning an author main entry.

In contrast to 3A and 3B2, however, even partial characterization of the condition of collective title was impossible to achieve. This was similarly the case with the comparable ALA rule 5, although there were no proof slips in the ALA sample with statements of the type noted above. In cases such as:

> A treasury of favorite bed-time stories
> Compiled[8] by
> Erwin Small

which were common in both samples, successful entry was accomplished without characterization of this condition or others in the rule that defined the nature of the contributions and the circumstances of their inclusion in the publication. In those such as the entry affected, successful entry was clearly dependent on the capacities of a cataloger.

---

[8] Or "edited," as in the case of many proof slips.

In contrast to the proof slips reported in Table 32, those noted in Table 33 and discussed above could not be assigned correct main entries by the models.

Summary. — In summary, the results of the application of the AA and ALA experimental rules to the recognitions on the proof slips in their respective random samples indicated that both sets of rules were highly successful and performed equally well - at the 97 per cent level. In contrast to the recognition rules, the selection rules in the models were similar, but not identical. Differences in the rules afforded a basis for evaluating their performance to determine if there was justification for choosing one model of entry selection over another.

Examination of the models for their ability to assign the required types of main entry (author or title) showed only slight differences. In general, both models had difficulty assigning correct title main entries. Similarly, there were only slight differences in the models with respect to the performance of their individual rules. Of the 9 AA and 11 ALA rules, all but four of the former and six of the latter assigned correct main entries to all their proof slips. The remaining rules in the two models assigned incorrect as well as correct main entries. Furthermore, the pairs of similarly constructed rules in the models, of which there were nine, were found to be comparable in both performance and frequency of use.

To determine the number of verbal rules of entry successfully characterized by the models, the proof slips assigned main entries by the experimental selection rules were examined for application of these rules. The results showed the ALA model to be considerably more successful on this point, with 22 or 61 per cent of the verbal ALA rules successfully characterized, in contrast to 8 or 42 per cent of the verbal AA rules. However, when the verbal rules of entry in each corpus were correlated according to their comparability determined in the study, the evidence indicated that an almost equal number were successfully characterized by the two models.

Examination of the incorrect entries in both samples showed that failure in the characterization of the verbal rules was attributable to conditions of entry that were either conflicting or too obscure to permit characterization in the experimental rules. In the case of both models, a small and almost equal number of verbal rules was involved. These rules, in turn, accounted for only a small percentage of the incorrect entries in both samples. The rest resulted from unsuccessful recognitions produced from application of the recognition rules and from incomplete characterization; namely, a failure to examine or provide for certain environments associated with patterns of authorship in the verbal rules. With alterations in the recognition rules and the selection rules, the majority of the proof slips with entry errors attributable to these latter sources could be assigned correct main entries by the models.

3. Complexity as related to the trees of the rules. — In addition to the accuracy of the experimental rules, it was proposed that the tree diagrams constituting the experimental selection rules in the AA model would be less complex than those in the ALA model. This, in turn, was because of certain factors related to the number and composition of their verbal counterparts in the two cataloging codes.

As applied to the models, complexity in the characterization was identified with the number of rules, nodes in the trees of the rules, and the steps necessary to achieve entry under the rules. Although these measurements alone appeared trivial, from the viewpoint of model construction and use, they assumed an especially important factor in choosing between the models. If, for instance, as the findings above suggest, the models are comparable in performance and in their characterization of verbal rules of entry determined comparable in the study, then surely the model whose characterization achieves this with a minimum of rules, specifications, and steps is preferable.

Measurement of the models involved a simple counting procedure, the results of which are reported in Table 34. Contrary to the prediction, there was very little difference in the models with respect to these measures. Although there was some support to indicate that the characterization in the AA model was simpler, i.e. more economical in the number of specifications at the nodes, etc., it was hardly substantial for justifying a choice of the model.

Table 34

Complexity in the Models as Measured by Number of Rules, Nodes, and Steps for Selecting Main Entry

Models	Rules	Nodes	Average Number of Nodes per Rule	Steps	Average Number of Steps per Rule
AA	9	66	7.3	112	1.2
ALA	11	90	8.2	249	2.3

Table 35

Redundancy in the Verbal Rules as Observed by the Number of Reduced Rules in the Models

Corpus	Number of Rules	Number of Redundant Sets	Average Number of Redundant Sets per Rule	Number of Reduced Rules in Models
AA	19	21	1.1	11
ALA	36	23	.63	9

There are, of course, other measures of complexity that are employed for evaluating models, such as amount of computation, description space, etc., although these are relevant mainly to the performance or actions of computer programs. Similarly, other criteria, such as cost and time in handling identical problems or data, are frequently employed in the evaluation. The application of these latter two especially could lead to a different interpretation of the data in Table 34. When interpreted in terms of time and cost, for instance, the economy achieved in the characterization in the AA model might be substantial to support a choice of that model. However, due to the scope and design of the study, these criteria could not be applied.

4. Causal factors as related to the characterization and performance of the models. —— The causal factors, which were said to influence the performance of the models, related to the number and composition of each corpus of main entry rules. In particular, it was specified that variables of size, complexity, redundancy, and specificity would determine the amenability of each corpus to formal characterization and, in turn, affect the complexity of the characterization, and its ability to assign correct main entries. Operational definitions of these variables were formulated, which were applied to the two sets of main entry rules (see Chapter III, pp. 55-57).

Since formulation of the definitions was necessarily arbitrary, it was recognized that their justification, in fact, would have to be made a posteriori, that is, in view of the adequacy of the characterization. For example, it was pointed out that a corpus of rules could be determined "complex" in terms of the quantitative measures specified by the definition for that variable, but unless these were observed in the characterization in some determinable way (ratio of nodes, steps in the rules, etc.), then the operational definition posited for that variable, or even the variable itself would not be justified as a relevant factor of amenability.

The operational definitions of the variables and the findings resulting from their application to each corpus of main entry rules (see Chapter III, pp. 75-101) were examined in terms of the characterization developed in the study and the test of accuracy performed by the investigator. The results of this examination are discussed under the subheadings for the four variables.

Size. —— Size or the number of rules of entry may (although it was not found to) be a relevant factor of amenability in the study. Testing of the hypothesis showed that there were almost twice as many comparable rules of entry in the ALA corpus, 36 as compared to 19 in the AA corpus, to be characterized. However, in the characterization, there were only two additional rules in the ALA model, 11 in contrast to 9 in the AA model. Thus, while differences in the two sets of main entry rules on this point were appreciable, these were not found or reflected in the two characterizations developed in the study.

Possibly, with expansion of each corpus, that is, the addition of more AA and comparable ALA rules of entry, differences in the characterizations might be more pronounced. However, this assumes that there will be a multiplicity of rules in the ALA corpus that will be special rules, requiring an equal or proportionate number of separate, special rules in the characterization. Although most of the rules in the ALA corpus that were characterized in the study were "special" rules, in the sense that they were designed for specific cases (type of publication, for instance, rather than conditions of authorship), these required, as noted above, only two extra rules in the characterization.

Redundancy. —— While redundancy or unnecessary duplication of rules was considered a factor of amenability, the quantitative measure posited for this variable was inadequate for determining its presence in the rules. Testing of the hypothesis showed that there was considerably more redundancy in the AA corpus than in the ALA corpus. Using the structural descriptions to determine equivalency in the rules, 21 sets of rules in the former corpus were found redundant. This was an average of 1.1 sets per rule. In contrast, 23 sets, or an average of .63 sets per rule, were found redundant in the latter corpus. However, these differences were reversed in the characterization.

In the characterization, the variable was observed in the number of reduced rules comprising each model. Table 35, which combines the findings on the variable, shows that the ALA corpus of 36 rules, with approximately 63 per cent found redundant, was characterized in 11, or about one third the number of rules in the corpus. The AA corpus of

19 rules, on the other hand, with over 100 per cent redundancy, was characterized or reduced to only 9 rules, or about one half the number of rules in the corpus. In short, the characterization of the AA corpus showed that there were considerably fewer redundant rules in the corpus than the structural descriptions had indicated.

As earlier examination of the structural descriptions of the AA corpus showed (see Chapter III, p. 76), the construct and verbal import of rule 19B made it redundant with every other rule in the corpus. Thus, it was found that the redundancy in this rule alone accounted for 18 of the 21 sets of rules that were determined structurally equivalent in the corpus. However, development of the characterization showed this rule to be a "recycling" rule (a direction to begin the entry process again using the previous rules), rather than a redundant rule (an unnecessary repetition of the features and specifications of entry in the previous rules).

This finding largely explains the discrepancies in the results noted above. On the basis of this finding, only three sets, or an average of .15 sets per rule were redundant in the AA corpus, a considerably smaller percentage of redundancy than was found in the ALA corpus. This smaller percentage, in turn, would similarly account for the considerably smaller number of reduced rules in the AA characterization.

Complexity. — Complexity, which was measured by the number of conditions specified in each corpus of entry rules, was found to be a relevant factor of amenability in the study. Testing of the hypothesis showed that there were 53, or an average of 2.7 conditions per rule in the AA corpus, which compared with 91, or an average of 2.5 conditions per rule in the ALA corpus.

The variable was observed in the characterization by the number of nodes in the trees in the experimental rules. Insofar as the characterizations were abstract representations of each corpus, there was no one-to-one correspondence, as such, of specifications at the nodes and verbal conditions in the rules. However, the slight differences in the number of conditions in each corpus were similarly found in the number of nodes in the characterizations. Table 36, which combines the findings on the variable, shows that there were 66, or an average of 7.3 nodes per rule in the AA model, as compared to 90 nodes, or an average of 8.2 nodes per rule in the ALA model.

Specificity. — As in the case of complexity, specificity, which was measured by the number of exact conditions in each of the entry rules, was found to be a relevant factor of amenability in the study. Bibliographic conditions in the rules were considered exact if they permitted objective observation, in contrast to subjective intuitive interpretation (for example, "principal responsibility" defined as the first named author on the title page, in contrast to the condition without any definition).

Testing of the hypothesis showed that both sets of entry rules were equally specific. Of the 53 conditions in the AA corpus, 37 or

Table 36

Complexity in the Verbal Rules as Observed by the Number of Nodes in the Experimental Rules

Corpus	Number of Rules	Number of Conditions	Average Number of Conditions per Rule	Number of Nodes	Average Number of Nodes per Rule
AA	19	53	2.7	66	7.3
ALA	36	91	2.5	90	8.2

Table 37

Specificity in the Verbal Rules as Observed by the Number of Conditions Amenable to Formal Characterization

	Exact Conditions		Inexact Conditions		Inexact Conditions Resulting in Incorrect Main Entries
Corpus	Rules of Entry	Experimental Rules	Rules of Entry	Experimental Rules	
AA	37	37	16	16	4
ALA	70	70	21	21	5

an average of 1.9 conditions per rule were determined exact. This compared with 70 of the 19 conditions, or an average of 1.9 conditions per rule, in the ALA corpus.

In the characterization, the variable was observed in the specifications contained at the nodes and roots of the trees of the experimental rules. In the case of each corpus, the bibliographic conditions of entry in the rules were characterized in the form of specifications ("Is $\mathcal{V}_1$ in larger type?" etc.). As was previously mentioned, there was no one-to-one correspondence of the specifications and verbal bibliographical conditions. Likewise, as the illustration above shows, the form of the specifications bore little resemblance to their verbal counterparts. This abstractness of the characterization consequently makes it difficult to observe which, as well as how many, of the conditions could be characterized in the study.

Observation of the variable in the characterization, however, is facilitated by examining the specifications at the nodes and roots of the trees and the bibliographical conditions for the two sets of entry rules given in Tables 13 and 14 (see Chapter III, pp. 91-94, 95-99). For example, taking the specification above, reference to Table 13 shows the specification is a characterization of AA conditions 7 and 38; Z at the root of ALA experimental rule 5, for instance, is a characterization of ALA conditions 47, 49, 55, 56, 72, 73, and 81 in Table 14, and so on.

The results of examining the specifications in the trees of both models showed that all the conditions determined exact in the tables were amenable to characterization; likewise, all those determined inexact were found impossible to characterize within the limitations of the study.

Table 37, which combines the findings, also shows the number of inexact conditions which, lacking characterization, caused the models' assignment of incorrect entries in the samples. It is of particular interest to note the small number of these conditions that caused the entry failures. This finding points up a feature of the characterizations, which was briefly noted in connection with their formulation (see Chapter IV, pp. 137-38), namely, that successful assignment of the main entry could be achieved despite non-characterization of a bibliographic condition.

For instance, although it is impossible to characterize the condition of "original author" in ALA 22A, correct main entries were assigned to proof slips requiring application of the verbal rule. In such cases, it was found that characterization of those conditions in the rule, that were directly observable on the title page, was adequate for the purposes of assigning successful entry choices.

Summary. — Of the four causal factors, which were said to determine the amenability of each corpus of main entry rules to formal characterization, only two, complexity and specificity, were found to be relevant factors in the study. Earlier measurement of these two variables in the verbal entry rules (see Chapter III, pp. 79-100) showed that, contrary to the prediction of the hypothesis, both sets of rules were almost equally amenable to formal characterization. In view of these findings, the study hypothesis would predict that the two characterizations would be almost equally able to assign correct main entries. The findings, resulting from the test of accuracy performed by the investigator, fully supported this prediction. Both models performed at the same level of accuracy. Further examination of the models presented in sections 1 to 3 in the chapter, showed the models to be comparable on a number of related variables. The results of these examinations are summarized at the end of each section.

In the following chapter, the adequacy of the models to accomplish entry selection automatically, that is, objectively and non-intuitively, will be discussed, and the findings resulting from an experiment involving their application to a problem set of title pages will be presented.

CHAPTER VI

SECOND TEST OF ADEQUACY

Chapter V presented and discussed the findings relating to the adequacy of the simulation models to generate correct main entry choices. This chapter will discuss the second test, which demonstrated the adequacy of the models to accomplish selection of entry automatically, that is, objectively and non-intuitively. This test was performed in an experiment, which involved the application of the models by two groups of participants to a specially selected problem set of English language monographs.

This chapter follows the same organization and presentation of material as that used in Chapter V. The first section will discuss the procedures used in the performance of the test, with attention to the design of the experiment, the selection of the title pages, and groups of participants involved. This will be followed by a second and final section, in which the findings resulting from these procedures are presented.

Procedures

It was the objective of the second test of adequacy to demonstrate that selection of entry can be accomplished in a completely objective, mechanical manner without the use of human intuition and foreknowledge. The necessity for conducting this test separately was mentioned in the previous chapter (see pp. 167-68). Although it was the original intention of the investigator to conduct a single test, that would demonstrate both the mechanicability and accuracy of the models, the difficulties relating to the acquisition and preparation of needed materials showed this to be unfeasible. By conducting two separate tests, it was found that these difficulties were greatly diminished.

The test of mechanicability, as it was conducted in the study, required direct application of the models to the author and title statements on the title pages of books. To demonstrate the mechanicability of this process, a group of participants, totally lacking in professional library training and experience, was asked to apply the models to a selected set of English language title pages. In an attempt to determine the effectiveness of the models insofar as they assist professionally trained librarians to assign main entry choices mechanically, a second group of participants, with professional library training and cataloging experience, was formed. This group was assigned the same task as the first group, using the same set of English language title pages.

In the following pages, the test of mechanicability[1] is described in detail. The procedures used to conduct the test are discussed first, with sections on the selection of title pages, the formation of the two groups of participants, and the adminstration and scoring of the test.

---

[1] "Mechanicability" is used as a convenient descriptive phrase throughout the chapter, referring specifically to the manual application of the models. Insofar as development of the models utilized alphabetical and numerical devices (letters, letters combined into word units, numbers, and binary features of a language, and not the meanings or semantic connotations of a language), application was seen to optimize objective, mechanical responses on the part of the participants.

Selection of title pages. --- A problem set of forty-four title pages, comprising two subsets of twenty-two pages each, was used in testing the models for their adequacy to produce main entry selections mechanically. Although application of the models to the same set of title pages was perhaps most desirable, a number of considerations indicated that this approach was inadvisable.

With one exception, the recognition rules in the models were identical, while the selection rules, although not identical, were similar in form and construction. Consequently, it was believed that the application of two closely related models to one set of title pages would result in a familiarity with both the models and the sample which would lessen the mechanicability of the process. An element of confusion was also seen to result in the case of some title pages, where application of the models would produce conflicting types of entry. Remembrance of the assignment of an author main entry for a title page involving the application of the AA model, for instance, was seen to result in possible misapplication of the ALA model, where the entry assignment would be under title. To counteract these biases, a lapse or interval of time between applications was necessary. However, practical considerations with respect to the participants' involvement, and the time limitations of the study, made this approach unfeasible.

In using the two subsets of title pages, objective measurement required that they be matched for identical problem types. Initially, a title page was selected for each rule in the AA corpus and its counterpart in the ALA corpus. However, in the case of certain rules, mainly the redundant rules in each corpus, the patterns of authorship on the title pages were identical. Consequently, a single title page was sufficient to illustrate the general application of these rules. On the other hand, in the case of rules such as AA 1A and 3B1 and their counterparts, several unique patterns of authorship were associated with each rule. Representative illustration of these rules and application of the models called for several title pages.

Final selection of title pages was based on these considerations. However, the decision as to the size of the sample set was dictated by practical considerations with respect to the preparation of materials and the time required of the participants to apply them. Practical application required that there be a set of recognition and selection rules for each title page selected for testing. Thus, for participants in a group to apply the models simultaneously, six sets of rules were needed for each title page. Pretesting with trial title pages also indicated that participants could be expected to spend six to eight minutes to complete application of a set of rules to a title page. In view of these considerations, the sample size of forty-four title pages was reached, with the expectation that each participant would complete application of the models to the title pages in approximately four to six hours.

Table 38 shows the selection of title pages in terms of their application of AA and ALA rules determined comparable in the study. Together, the forty-four title pages represented the application of fifteen sets of verbal rules, or approximately 79 per cent of the sets characterized in the study. This accounted for fifteen AA rules and eighteen comparable ALA rules. The discrepancies between the number of title pages noted in the columns and the totals given at the bottom of the table result from the fact that, in the case of three title pages in each subset, two verbal rules had to be applied to assign the main entry. The number of title pages representing application of ALA rules 4A(2) and 5A(1) are given twice, which accounts for the additional discrepancy in the ALA set. It will be noted that these rules were considered comparable to several AA rules.

To facilitate analysis and comparison, each subset of title pages was arranged and identified alphabetically by a lower or upper case letter. The subset involving application of the ALA model was identified by upper case or capital letters: A to V. The subset involving application of the AA model was identified by lower case letters: a to v (see Appendix E, pp. 415-60). Arrangement or sequence of title pages was determined by a desire to avoid selection of main entries under successive model rules. Consideration was also given to placing title pages containing simpler patterns of authorship (a single author name and recognition marker, for instance) throughout the subsets to prevent the impression of progressive difficulty in rule application.

Table 38

Sample of Title Pages Correlated by Corpus of AA and Comparable ALA Rules of Entry

AA Sample	AA Rules of Entry	ALA Sample	ALA Rules of Entry
5	1A	5	2
1	2C	1	32
1	3A	1	3A
		1	3D
3	3B1	1	4A(2)
		1	4B(1)
		1	6B(1)
1	3B2	1	4A(2)
2	4A	2	5A(1)
		1	5B
2	5A	2	5A(1)
		1	5A(1a)
1	5B	1	5A(2)
1	7A	1	23A
2	8B	2	19E
1	9A	1	6D(2)
2	14A	2	20
1	15A	1	21A
1	16A	1	3E
1	19B	1	5C(4)
Total 22	15	22	18

Finally, as in the case of the previous samples tested in the study, the shipping and serial numbers of the Library of Congress cards for the items were used as an indicator of the cataloging code employed to assign the main entries. In the subset used to test the mechanicability of the ALA model, cards for the items had printing dates of (19) 49, 53, 55, 58, 60, 61, 64, and 65, and were from shipments received at the University of Illinois and local area public libraries prior to 1966. The items in this subset were considered to be cataloged unquestionably under ALA rules of entry. In the subset used to test the AA model, cards for the items had serial numbers (19) 68, 69, 70, 73, 75, 76, and 77, and were from shipments received at the University of Illinois and local public libraries in the period 1968-1969. As previously mentioned, because of the Library of Congress' policy of superimposition, the dates in these cases were only partial assurance that the AA code had been used to assign the main entries. In a few cases, the choice of main entry confirmed use of the code; otherwise, complete confirmation was lacking.

Selection of participants. — Two groups of six persons each were formed to apply the models to the sample of title pages. The first group was composed of persons totally lacking in library training and experience, who were to function as automata in the entry selection process. On the basis of only an introduction to the model rules and tree structures, participants in the group were to demonstrate that selection of entry can be accomplished in a completely objective, mechanical way.

With the help of the Student Employment Service at the University of Illinois, six potential candidates, all freshmen students, were identified for inclusion in the group. The candidates were interviewed individually and two were rejected. In one case, it was found that the student had worked briefly as an aide in a high school library. In the other case, the student, through personal association with a professional librarian, was knowledgeable and generally aware of library procedures and, in particular, cataloging functions and tools. These candidates were replaced with two other freshmen. With the exception of one freshman, all were first semester students at the University of Illinois. Three were enrolled in the College of Liberal Arts and Sciences, two in the College of Commerce, and one in the College of Engineering.

The second group of six persons was selected to test the effectiveness of the models insofar as they would assist professionally trained librarians to formulate entry choices mechanically. This group was composed of five employed members and one retired member of the cataloging staff at the University of Illinois. All members held fifth-year library degrees, and had an average of ten years' professional cataloging experience. For the most part, cataloging experience of the members was restricted to the use of the ALA code. However, three indicated having limited use of the AA code, while the others admitted to having personally familiarized themselves with parts of the AA code and, in some cases, having analyzed and compared the code's general rules of entry with those of the ALA code. The three members with experience in using both codes were responsible for their own original cataloging. The retired member had been a senior reviser in the cataloging department. The two remaining members did general cataloging under the immediate supervision of a reviser.

Administration of the test. — Testing of the models was conducted separately for each group. Since there were some differences in procedures, administration of the test is described separately for each group.

Testing of the models performed by the first group of participants, comprised of freshmen students, began by issuing each participant two sets of instructions for the two sets of rules comprising the models, and a sample title page illustrating their application (see Appendix D, pp. 375-414). Instructions consisted of two parts: a brief introduction, stating the purpose and function of the rules, followed by a discussion of points relating specifically to their application to the sample title page. The sample title page was accompanied by a copy of the ALA recognition and selection rules, illustrating points in the application of the two sets.

Each participant was given two days to read the introductory material, followed by a personal meeting with the investigator, at which time a subset of title pages was issued. Prior to issuing the material, the participant was given time to direct questions and comments to the investigator. These resulted in briefly reviewing the points in rule application reiterated in the instructions and illustrated in their application to the sample title page.

Each title page, which was xeroxed in multiple copies, was stapled in a manila folder, accompanied by a set of the model's recognition and selection rules, arranged and identified in the manner described earlier. With each title page, the participant applied the recognition rules, and selection rules in that order, using the respective spaces and blocks to record the steps and decisions in applying each set. The first subset of title pages to be issued involved application of the AA model, followed by the second subset involving application of the ALA model. Time spent was considered of general interest. Upon completion of each subset, the amount was recorded on a time card for the participant.

In view of the expected duration of testing, four to six hours, the investigator first considered administering the test to the group as a whole over a period of two to three days. However, this plan was discarded for several reasons. Schedules and commitments were so varied that only two participants could meet at any one time. Accommodation of all participants necessitated administering the test over a period of two to three weeks at brief and awkward times of day. Moreover, such prolonged testing with the group was seen to provide opportunities for the participants to discuss the material and their performance with one another. Consequently, three two-hour meetings were tentatively arranged with each participant to apply the models, with the understanding that testing would cease upon completion of both subsets of title pages. At each meeting a participant completed as many title pages in a subset as time allowed. At the close of the meeting, the completed title pages as well as any remaining in the subsets were submitted to the investigator. Testing was conducted in this manner for each participant during December 1969.

The same preliminaries were followed in administering the test to the second group of participants, comprised of catalogers. Each participant was issued a set of instructions and a sample title page two days prior to meeting with the investigator. However, a group meeting was held with all the participants, where the investigator reviewed the material and issued the first subset of title pages and accompanying sets of rules. Since personal supervision of the participants could not be satisfactorily arranged due to work schedules and personal commitments, it was decided that the participants would perform the test independently. The group meeting then allowed the investigator to issue the test material and, at the same time, to elaborate more extensively on the objectives of the test, and the precautions to be observed in ensuring their attainment.

Although it was realized that discussion and comparison of results could not be controlled, participants were instructed to reserve their comments until completion of the test, and to direct any questions regarding procedures and results to the investigator. With easy access to library catalogs, National Union Catalog records, etc., it was also realized that the main entries for the title pages could be easily obtained or verified. However, participants were not informed or assured that the models would permit successful recognition of an author name or agreement in the selection of the main entry type or heading.

Administration of the test was conducted over a two week period in February 1970. In the first week, participants completed the subset of title pages involving application of the AA model. These were submitted to the investigator, and the subset involving application of the ALA model was issued and completed in the second week. Time spent in applying the models was recorded by the participants on a time card issued with each subset.

In view of the exploratory nature of the study, the investigator thought it desirable to obtain the reactions of the participants. In this connection, a brief questionnaire was issued to the participants upon completion of the testing (see Appendix F, pp. 460-62). Question 1, relating the the sequence or order of title pages followed in model application, was directed only to the catalogers, who performed the test independently. The three remaining questions were directed to all participants. Since it was not possible to anticipate the nature and kinds of responses the participants would record in applying the models, the investigator arranged to interview each participant after obtaining the test results. If there were difficulties in applying the models, it was believed that the interviews would be a particularly useful device for eliciting information pertinent to the experiment and to the study in general.

Scoring. — The criterion of mechanicability used for testing the models was the correct application of the recognition and selection rules comprising each model. Application of the rules was determined correct if the instructions and actions ordered by the rules were followed and completed successfully for each title page in the sample set. Correct application of a model's rules to all twenty-two title pages in a subset was an indication of the model's high degree of mechanicability. In contrast, incorrect application of the rules to all the title pages was an indication of the model's low degree or absence of mechanicability.

While correct application of the rules called for the assignment of correct main entries, the latter alone was not considered to be indicative of a model's mechanicability. Incorrect application of the rules, for instance, might still permit the assignment of a correct entry choice, in which case the entry results would be accidental and reflect circumstances of the data, etc., rather than the mechanicability of the rules. For example, it might be found that the instructions for recognizing the marker AND could not be followed. In cases involving the recognition of two or three author names, as in title pages AAh and ALAH (see Appendix F, p. 423, 445), where the main entry is under the first author's name, the selection rules in the models would permit the assignmemt of the correct entry choice.[2] However, the entry results in these cases would not account for nor illustrate the model's lack of mechanicability.

Since limitations of the study ruled out considerations of the models' technical feasibility, it was realized that the employment of humans to perform the test was a severe restriction. Although the first group of participants lacked library training and experience and was selected to function as automata in the test, rigor and objective criteria could not control their conscious thought, feeling, or intuition. Consequently, their performance of the test was considered to be an approximation only of simulation by an automatic device or mechanism. In the case of the second group of participants, comprised of professional librarians, their cataloging knowledge and experience made unbiased application even more difficult. For this reason, the test was considered to be a measure only of the effectiveness of the models to assist these persons to formulate main entry choices mechanically.

For scoring purposes, the results of each participant's application of a set of rules were compounded into a single score for the group. Thus, for each group, there was one score for application of the recognition rules, and another for application of the selection rules in each model. In an attempt to determine the nature and extent of misapplication, the rules and title pages were identified and examined for their frequency and source of entry error. Since time to apply the models was considered to be a point of general interest, each group was also scored in terms of its speed in applying each model to the respective subset of title pages.

---

[2] Assignment of the main entry would be accomplished under selection rules 8 and 10 in the AA and ALA models respectively. However, in cases involving the recognition of four or more author names, as in title pages AA j and ALA J, where main entry is under title, inability to apply AND would result in the assignment of incorrect main entries by the models' selection rules.

Pretesting of the models and procedures was done in several stages, using trial title pages tested on both professional catalogers and freshmen students. A detailed discussion of this topic is presented in Chapter V (see pp. 182-83) in connection with the first test of adequacy performed on the models.

Findings

In this section, the findings resulting from the test of mechanicability performed by the two groups of participants are presented. Inasmuch as possible, the presentation follows the arrangement in presenting the findings resulting from the first test of adequacy. The application of the models is discussed under four numbered subheadings. In the first subheading, the findings related to the application of the recognition rules in the models are presented. In the second subheading, those related to the application of the selection rules are presented. These are followed by a third subheading, in which the models are compared briefly for speed in application by the two groups of participants. In the final subheading, the results of the post-test questionnaire and interviews are briefly reviewed.

1. Mechanicability, as related to the application of the recognition rules. — Correct application of the recognition rules in the models called for the identification of the author and title statements on the title pages of the sample set, and the recognition of the markers BY, AND, WITH, OF, BETWEEN, FROM, 'S,S' and accompanying data in these statements, as required. In the test conducted in the study, the number of title pages, whose data had been correctly identified, indicated that the recognition rules in the AA and ALA models were mechanical at the 91 and 95 per cent level respectively for the group comprised of freshmen students (identified as Group 1), and at the 82 and 77 per cent level respectively for the group of catalogers (identified as Group 2). The results of their application to the paired subsets of title pages are presented in Table 39.

Since all but one of the recognition rules in the two models were identical, the comparability in their application was anticipated. As the table shows, this was the case for both groups, with Group 1 successfully applying the ALA rules to 21 title pages, and the AA rules to 20 title pages in the two subsets, and with Group 2 successfully applying these rules to 18 and 17 title pages respectively. However, application of the recognition rules by the two groups differed by approximately 9 per cent in the case of the AA set and 18 per cent in the case of the ALA set. In the case of Group 1, designated as automata in the test, the results showed the two sets of rules to be highly, as well as almost equally, mechanical. Application of the rules

Table 39

Number of Title Pages in Each Subset with Correct Application of Recognition Rules

	Recognition Rules					
	AA			ALA		
	Title Pages	Correct	Percentage	Title Pages	Correct	Percentage
Group 1	22	20	91	22	21	95
Group 2	22	18	82	22	17	77

by Group 2 resulted in the lower scores of mechanicability noted in the table. In view of the findings, the results of the latter strongly suggested that mechanical application of the rules was difficult for experienced catalogers to accomplish.

Sources of incorrect application. — Examination of the rules and title page data to which they were incorrectly applied was considered of particular importance in determining the nature and source of their misapplication and, in turn, the limitations as well as the potential of their mechanicability. The findings resulting from this examination are presented in Tables 40 to 43.

Examination of the title pages to which the recognition rules were incorrectly applied revealed that with the exception of one title page in each subset, the two groups misapplied the rules to different title pages. Similarly, with the one exception noted, which involved misapplication of the same recognition rule in each model, the two groups misapplied different recognition rules. The results of this examination are reported in Tables 40 and 41 for the AA and ALA rules repectively.

In the case of the AA rules, only 2b, which called for recognizing the marker 'S,S' in the environment of drawings, foreword, etc., was misapplied by both groups. Misapplication in both cases involved title page g. In the case of the ALA rules, only 2c, which called for recognizing 'S,S' without BY in the substring, was misapplied by the two groups. Misapplication involved title page V. Although both rules

Table 40

Title Pages and AA Recognition Rules Incorrectly Applied

	Title Page	Frequency	AA Recognition Rule
Group 1	c	1	1b
	q	1	2b
Total	2		2
Group 2	j	3	6b
	k	3	6a
	q	1	2b
	u	1	6c
Total	4		3

Table 41

Title Pages and ALA Recognition Rules Incorrectly Applied

	Title Page	Frequency	ALA Recognition Rule
Group 1	V	3	2c
Total	1		1
Group 2	E	1	5
	L	4	6a
	P	1	6a
	Q	1	1a
	V	2	2c
Total	5		4

involved recognition of the same marker, it is difficult to assign the groups' misapplication to a lack of mechanicability in the models. The low frequency with which the two rules were applied to the title pages (only one participant in each group in the case of AA 2b, and half or less of the participants in the groups in the case of ALA 2c) did not support the conclusion. Moreover, both rules had identical counterparts in the models, that is, AA 2b had an identical counterpart in ALA 2b, and ALA 2c had its counterpart in AA 2c. However, in each case, only the rule in one model was misapplied. In the case of the other title pages and rules noted in the tables, the differences in the groups (with respect to the title pages and the rules that were misapplied) gave little indication that misapplication of either model was due to its lack of mechanicability.

On the contrary, examination of the title page data and the recognitions recorded for these rules revealed that human error and oversight, rather than mechanicability, were the major source of their misapplication. The rules and types of error accounting for their misapplication are presented in Tables 42 and 43 for the AA and ALA sets respectively .

Incorrect abstraction of the strings (initial and sub), and the misidentification of markers and their environments (recording a marker's presence, although physically absent in the title page data, or recording a marker's environment, although physically absent in the rule) accounted for the misapplication of the rules by Group 1.

Table 42

Types of Errors Accounting for Misapplication of AA Recognition Rules

	Title Page	AA Recognition Rule	Type of Error
Group 1	c	1b	abstraction of strings
	q	2b	misidentification of marker
Group 2	j	6b	omission of marker recognition
	k	6a	omission of marker recognition
	q	2b	misidentification of marker
	u	6c	omission of marker recognition

Table 43

Types of Errors Accounting for Misapplication of ALA Recognition Rules

	Title Page	ALA Recognition Rule	Type of Error
Group 1	V	2c	misidentification of marker
Group 2	E	5	partial marker recognition
	L	6a	omission of marker recognition
	P	6a	omission of marker recognition
	Q	1a	abstraction of strings
	V	2c	misidentification of marker

Although these errors were also made by Group 2, the tables show that failure to recognize the markers (their physical or implied display in the title page data) largely accounted for misuse of the rules by Group 2. This latter error, in particular, was of considerable interest for what it revealed about the composition of the group. The errors of both groups are discussed briefly below.

Misidentification of the markers and their environments was made by several participants in each group, and involved the same title pages, g and V (for reference to these and other title pages discussed in the section, see Appendix E, pp. 415-59; for reference to the recognition rules, see Appendix B, pp. 328-35, 349-55). In the case of g, a participant in each group applied AA rule 2b to produce the recognition: #Playboy's# #selected by the editors#, when there was no provision in the rule for recognizing the marker 'S in this environment. In the case of V, half of the participants in Group 1 and two in Group 2 applied ALA rule 2c to produce the recognition: #THE AMERICAN PEOPLE'S#, when there was no physical display of the marker 'S in the title page data. In both cases, the errors unquestionably resulted from misreading either the data or the rules.

Errors with respect to the omission of marker recognition were made only by the catalogers in Group 2, and involved three title pages in each subset: j, k, and u, and E, L, and P. Omission of marker recognition occurred in connection with the markers AND and OF in AA and ALA rules 6 (identical in both sets) and ALA 5, respectively.

With the exception of title page *E*, application of the rules for these markers resulted in the recognition of non-author names, which could not be avoided with the given construct of the rules. Although unintentional, the required recognition of these non-author names indirectly provided a measure of the group's objectivity with respect to applying the rules — and approximating machine simulation of the recognition process. The correct application of these rules by all the participants in Group 1, to a large degree, was assurance of the group's overall objectivity. In contrast, there were only two participants in Group 2 who correctly applied the rules to the six title pages. While human oversight was a possible explanation for omission of the marker recognitions involved, post-test interviews revealed that the four participants in this group were reluctant to recognize the non-author names, which resulted from applying these rules. For instance, in the case of title pages *k*, *L*, and *P*, where the substrings on the pages had the presence of the marker AND implied at the ends of their strings, application of 6a required the respective recognitions (indicated by underscoring):

(k) #... by Eleanor Terry Lincoln *[AND] Smith College#*

(L) #... BY LASZIO *[AND] MATULAY#*

(P) # By Floyd C. Bard *[AND] as told to [AND]*
Agnes Wright Spring#

Recognition of the marker in the middle of the author name, as in *L*, was found to be particularly disturbing, as was the case with title page *j*, where the substring had the marker's presence implied in the form of a comma, resulting in:

#... Robert G. Dixon [AND] Jr.#

The reluctance to recognize these incongruities clearly pointed up the difficulty, in general, that this group experienced in applying the rules objectively.

Of the three types of errors, those with respect to abstracting the title page into the two strings (initial and sub) were the least frequent. Only one participant in each group made this error, which involved two different title pages, *c* and *Q*. In the case of *c*, a participant in Group 1 applied AA 1b to produce the recognition of a substring: #[BY] Book #, where application of the rule was not required, while in the case of *Q*, a participant in Group 2 applied ALA 1a to produce the recognition: #*by the Experts BY THE EDITORS*# as the substring in place of: #BY THE EDITORS#. In the latter case, the abstraction was triggered by the appearance of a BY in the data, where the participant overlooked the physical break that marked its presence in two statements - and, hence, two strings. In the case of *c*, however, it was found that the physical break in the data: Malcolm Boyd's/ Book of Days (the break indicated by the slash), caused the resulting abstraction. Lacking precise specifications as to the size of this break, the error was understandable. Unlike previous errors, it was a technicality that stemmed from the rules rather than human oversight.

Thus, with provision of the technicality and the exclusion of human errors noted, the scores attained by Group 1 would be increased by 9 and 5 per cent for the AA and ALA rules respectively, raising the mechanicability of both sets of rules to the 100 per cent level. In the case of Group 2, the scores would be increased considerably. However, the human errors made by this group were an important determinant of the rules' effectiveness in assisting catalogers to perform the entry selection process. Excluding the errors was irrelevant for determining the potential effectiveness of the rules for these people.

Summary of the recognition rules. –– The results of testing the mechanicability of the AA and ALA recognition rules on a selected sample of forty-four title pages by two groups of participants indicated that the rules in the models were mechanical at the 90 and 95 per cent level respectively for Group 1, acting as automata in the study, and 82 and 77 per cent respectively for Group 2, comprised of experienced catalogers. Since six of the seven recognition rules in the two models were identical, comparability in their application was expected for each group.

With one exception, the two groups misapplied the rules to different title pages. Similarly, with the exception noted, which involved misapplication of the same recognition rule in each model, the two groups misapplied different recognition rules. In the case of both groups, human error and oversight, rather than the mechanicability of the rules, accounted for the majority of the recognition errors.

Three categories of errors were identified: (1) misidentification of the recognition markers and their environments, (2) omission of marker recognition, and (3) incorrect abstraction of the two strings (initial and sub).

Errors in the first category were made by several participants in both groups and involved the same two title pages. In both cases, the errors unquestionably resulted from mis-reading, either the data or the rules.

Errors in the second category were made only by the catelogers in Group 2, and largely accounted for the lower scores attained by this group. While human oversight was a possible explanation for the group's failure to recognize the markers on the six title pages involved, post-test interviews with the participants revealed their reluctance to recognize the non-author names, which resulted from applying the rules.

Errors in the last category were made by only one participant in each group, and involved two different title pages. Human oversight accounted for the error in one case, while a technicality in one of the rules accounted for the error in the other case. Unlike previous errors, the latter error was the only one that stemmed from the rules.

With provision of the technicality and the exclusion of human errors, it was found that the scores attained by Group 1, designated an automaton in the test, would be increased by 9 and 5 per cent for the AA and ALA rules respectively, thus raising the mechanicability of both sets of rules to the 100 per cent level. In the case of Group 2, which was comprised of the catalogers, the scores would be increased considerably. However, the human errors made by this group were an important determinant of the rules' effectiveness in assisting catalogers to perform the entry selection process. Excluding the errors was irrelevant for determining the potential effectiveness of the rules for these people.

2. Mechanicability, as related to the application of the selection rules. —- Correct application of the selection rules in the models called for matching the structure of the recognitions recorded for the title pages (by ... and ..., etc.) with that contained at the roots of the trees in the selection rules (BY $\mathcal{X}$ AND $\mathcal{Y}$, etc.), and following the heuristics dictated by the match for the assignment of the appropriate main entry.

In the test conducted in the study, the number of title pages to which the experimental selection rules had been correctly applied, indicated that the rules in the two models were highly and almost equally mechanical. Application of the rules by Group 1 showed the AA and ALA models to be mechanical at the 91 and 95 per cent level respectively. Application by Group 2 showed the respective models to be mechanical at the 95 and 91 per cent level. The results of their application to the paired subsets of title pages are presented in Table 44.

Table 44

Number of Title Pages in Each Subset with Correct Application
of Experimental Selection Rules

	Selection Rules					
	AA			AIA		
	Title Pages	Correct	Percentage	Title Pages	Correct	Percentage
Group 1	22	20	91	22	21	95
Group 2	22	21	95	22	20	91

In contrast to the recognition rules, the groups differed by only 4 per cent in their application of the two sets of selection rules, with each group attaining a higher score in its application of one of the models. In this connection, the scores of Group 2 were of particular interest. In the case of both models, the scores of this group's application of the selection rules were noticeably higher than those for the application of the recognition rules. A comparison of the figures in the above table and those in Table 39 shows the group successfully applied the selection rules to an additional three title pages, or approximately 14 per cent of those in each subset. Although the cataloging knowledge and experience of the members in the group may have assisted them in their application of these rules, the results also suggested that the construct and heuristics of these rules, in particular, facilitated their mechanical application, a point discussed in greater detail later in the section.

Unlike the recognition rules, the selection rules in the models were similar but not identical. Differences in the selection rules with respect to their number, branching of trees, specifications at the nodes and endpoints, etc. were discussed in the previous chapter (see pp. 198-200). As in the test of accuracy, these differences were seen to provide a basis for evaluating the mechanicability of the rules to determine if there was some justification for choosing one model of entry selection over another. Both models were subsequently examined on a number of points. The findings which resulted from this examination are presented in Tables 45 - 49.

Application of the experimental selection rules. — Since the sample of title pages was necessarily small, there were for the most part only one or two title pages in a subset that required the application of a particular selection rule in a model. Consequently, although there were only a few title pages in the subsets to which the rules were misapplied, the recognitions of these title pages could well account for the groups' complete misapplication of several selection rules. The complete misapplication of a rule by all or the majority of the participants in the goups was seen to indicate the rule's lack of mechanicability. On the other hand, partial misapplication, where only a few of the participants misapplied the rule, or where the participants misapplied the rule to only some of the recognitions requiring the rule's application, was seen to suggest human errors or oversight rather than the rule's limited mechanicability. Thus, if it were found that the recognitions to which the AA model were misapplied involved only partial misapplication of the rules while those to which the ALA model was misapplied involved complete misapplication of the rules, then this difference would be supportive evidence of the superiority of the latter model.

To examine the models on this point, the rules of the models and the title pages to which they were applicable were listed for their application by the two groups of participants (for reference to the selection rules discussed in this section, see Appendix B pp. 336-48; 356-71). The results are reported in Table 45. As the columns of the

Table 45

Application of Experimental Selection Rules to Paired Subsets of Title Pages

AA Rules	Title Pages	Correct	Frequency		In-correct	Frequency		ALA Rules	Title Pages	Correct	Frequency		In-correct	Frequency	
			G_1	G_2		G_1	G_2				G_1	G_2		G_1	G_2
1	o	x	6	6				1	O	x	6	6			
2	c	x	5	6	x	1		2	C	x	6	6			
3	a	x	6	5	x		1	3	A	x	6	6			
4	r	x	6	6				4	R	x	6	6			
5	d	x	6	6				5	D	x	6	6			
	f	x	6	6					F	x	6	6			
	i	x	6	6					I	x	6	6			
6	j	x	6	6				6	E	x	6	6			
7	u	x	6	6					Q	x	6	5	x		1
8	b	x	6	6				8	J	x	6	6			
	e	x	6	6				9	U	x	6	6			
	g	x	6	6				10	B	x	6	6			
	h	x	6	6					E	x	6	6			
	k	x	6	6					G	x	6	6			
	l	x	6	6					H	x	6	6			
	m	x	6	6					K	x	6	6			
	n	x	6	6					L	x	6	6			
	p	x	6	6					M	x	6	6			
	q	x	5	6	x	1			N	x	6	6			
	s	x	6	6					P	x	6	6			
	t	x	6	6					S	x	6	6			
9	e	x	6	6					T	x	6	6			
	l	x	6	6				11	E	x	6	6			
	q	x	5	6	x	1			L	x	6	6			
	v	x	6	6					V	x	3	4	x	3	2

table show, all of the rules in the two models were correctly applied by half or more of the participants in the two groups. Of these, five or 56 per cent of the rules in the AA model and eight or 80 per cent of those used in the ALA model were correctly applied by all the participants to all the recognitions requiring their use. The remaining rules in the two models, four AA and two ALA, were partially misapplied. In the case of two AA rules, partial misapplication involved the same recognition.

It is of interest to note that, with the exception of ALA 11, both groups misapplied different rules in the models to different recognitions. Although there were several pairs of similarly constructed rules in the models, none of the misapplied rules were in pairs. Further, with one exception, none of the recognitions of the title pages involved were in pairs. These findings, along with the low frequencies of misapplication, gave support to the assumption that partial misapplication of the rules in the models was a result of human errors on the part of the participants. If this assumption is indeed correct, the evidence produced in the table would indicate that both models are equally and completely mechanical.

Source of incorrect application. — The six rules and the five title pages to which they were applicable were examined for the source of their misapplication. In the case of all six rules, the participants in both groups omitted application of the rules to the recognitions for the title pages. With one exception, omission of rule use was a result of human errors made by the participants either in matching the recognitions to the construct of the selection rules, or through previous misapplication of the recognition rules. In the case of the exception, omission of rule use was a result of a technicality in a recognition rule. The results of this examination are presented in Tables 46 and 47, for the AA and ALA rules respectively, and are briefly summarized in the following paragraphs under three letter subheadings.

A. *Misapplication of the recognition rules*. — All omissions of rule use in the ALA model were a result of the participants having misapplied the model's recognition rules. These involved two title pages, Q and V, the latter of which figured in the misapplication by both groups.

In the case of Q, failure to distinguish the physical break between the title and author statements on the title page resulted in the incorrect abstraction of the substring, producing the final recognition.

#by the Experts BY THE EDITORS OF SKI MAGAZINE#

The match of the recognition with the construct of ALA 10, Q BY ZZ, resulted in negation and the incorrect assignment of a title main entry under rule 11.

In the case of V, which had the highest frequency of misapplication (half of the participants in Group 1, and one-third of those in

Table 46

Source of Errors Accounting for Misapplication of AA Experimental Selection Rules

	Title Page	Applicable Rule	Rule Applied	Frequency	Source of Misapplication
Group 1	c	2	C*	1	technicality in recognition rule
	q	8,9	2*	1	mismatch of recognition and construct of selection rule
Group 2	a	3	2	1	mismatch of recognition and construct of selection rule

*Misapplication resulting in assignment of incorrect main entry.

Table 47

Source of Errors Accounting for Misapplication of ALA Experimental Selection Rules

	Title Page	Applicable Rule	Rule Applied	Frequency	Source of Misapplication
Group 1	V	11	2*	3	misapplication of recognition rule
Group 2	Q	6	10,11	1	misapplication of recognition rule
	V	11	2*	2	misapplication of recognition rule

*Misapplication resulting in assignment of incorrect main entry.

Group 2), the incorrect identification of marker 'S in the data resulted in the recognition:

#THE AMERICAN PEOPLE'S#

The match of the recognition with the construct of ALA 2, 𝔛'S,S' Z, resulted in the incorrect assignment of an author entry to: THE AMERICAN PEOPLE'S.

B. Mismatch of recognition and construct of selection rule. — Two of the three omissions of rule use in the AA model resulted from the participants (one in each group) mismatching the recognitions with the constructs of the selection rules. These involved title pages *a* and *q*, where in both cases, the recognitions were matched to the construct of AA rule 2, 𝔛'S,S' Z. Although rule 2 specifies that Z cannot be BY, the marker's physical presence in the recognitions:

(a) #Ed McMahon's#
#Drawings by Phil Interlandi#

(b) #PLAYBOY'S#
#selected by the editors of Playboy#

was overlooked.

In the case of *a*, application of the rule resulted in the correct assignment of an author main entry, while in *q*, the assignment of this entry proved incorrect. It is of interest to note that the recognition for the latter title page was incorrect due to the participant's misidentification of the environment for the marker 'S. However, despite this recognition error, a correct match with the constructs of rules 8 and 9 could have been accomplished, as it was in the case of another participant who made this recognition error. The match of incorrect recognitions with the constructs of correct selection rules occurred in the case of several title pages, as comparison of the present tables with Tables 40 and 41 shows. By either blocking or ignoring the recognition errors, the applicable selection rules succeeded in assigning the correct main entries, a point taken up in considerable detail in the previous chapter (see pp. 218-19).

C. Technicality in recognition rule. — Only in one case was omission of rule use a result of the rules rather than human errors. This involved title page c, where the recognition:

#MALCOLM BOYD'S#
[BY] Book#

was matched to the construct of AA rule 8: $\mathcal{Q}$ BY $\mathcal{Z}$Z, which resulted in assigning the author main entry to "Book." Although the abstraction of the substring was not called for, its recognition, as pointed out earlier, was understandable since there were no specifications in the recognition rules defining the size of the physical break (and hence, the division of the strings) between the title and author statements.

Thus, given these specifications and the exclusion of the human errors noted, the scores of the selection rules in both models would be increased. In the case of Group 1, designated an automaton in the test, the scores would be increased by 9 and 5 per cent for the AA and ALA rules respectively, raising the mechanicability of these sets, like those of the recognition rules, to the 100 per cent level.

An identical increase would be noted in the scores of Group 2. However, the human errors made by this group were an important determinant of the rules' effectiveness in assisting catalogers to perform the entry selection process. As in the case of the recognition rules, excluding the errors made in their application is irrelevant for determining their potential effectiveness for these people. In this connection, it is of some interest to note the differences in the scores for this group's application of the recognition rules and the selection rules. Although the higher scores for the selection rules could be partially explained by the familiarity of the process, the heuristics of these rules presented decidedly fewer frustrations in application than those developed for the recognition rules.

Summary of the selection rules. — In summary, the results of the application of the AA and ALA experimental selection rules to the recognitions for the title pages in the paired subsets indicated that the rules in the two models were highly and almost equally mechanical. Application of the rules by Group 1, designated an automaton in the test, showed the AA and ALA models to be mechanical at the 91 and 95 per cent level respectively. Application by Group 2, comprised of experienced catalogers, showed the respective models to be mechanical at the 95 and 91 per cent level.

Differences in the heuristics and specifications of the two sets of rules afforded a basis for evaluating their mechanicability to determine if there was justification for choosing one model of entry selection over another.

Examination of the models for the application of their individual rules showed that eight or 80 per cent of the rules used in the ALA model, in contrast to five or 56 per cent of those in the AA model, were correctly applied by all the participants to all the recognitions requiring their use. However, the remaining rules in both models were misapplied by only a few participants in the two groups, strongly suggesting that their misapplication was a result of human errors rather than limited mechanicability.

In examining the rules and title pages for their source of misapplication, it was found that with one exception misapplication was a result of human errors made by the participants, either in matching the recognitions of the title pages with the constructs of the selection rules, or through previous misapplication of the recognition rules. In the case of the exception, a technicality in one of the recognition rules accounted for the subsequent misapplication of a selection rule. In short, neither group was markedly superior to the other in its application of the models.

With provision for the technicality and exclusion of the human errors noted, it was found that the scores of Group 1 would be increased by 9 and 5 per cent for the AA and ALA rules respectively, raising the mechanicability of these sets, like those of the recognition rules, to the 100 per cent level. An identical increase would be noted in the scores of Group 2. However, the human errors made by this group were an important determinant of the rules' effectiveness in assisting catalogers to perform the entry selection process. As in the case of the recognition rules, excluding the errors made in their application was irrelevant for determining their potential effectiveness for these people. In this connection, it was noted that the group had conspicuously higher scores for application of the selection rules than for application of the recognition rules. Although the higher scores for the selection rules could be partially explained by the familiarity of the process, the heuristics of these rules decidedly presented fewer frustrations in application than those developed for the recognition rules.

3. Speed, as related to the application of the models. — Although it was not a factor in the application of the models, the time taken by the participants to apply the models was considered of general interest. Pretesting the models with trial title pages indicated that each participant could be expected to spend four to six hours applying the models to the sample set of title pages, or approximately five to eight minutes per title page. Examination of the time cards of the participants showed that the majority of the participants took a considerably longer amount of time to apply the models, ranging from six to thirteen and a half hours, or approximately eight to eighteen

minutes per title page. The results of this examination are reported in Table 48. In connection with the figures, it should be noted that the higher rates in the table indicate slower speeds in application.

As can be seen, differences between the groups as well as the models were small. Group 1 applied both models somewhat faster than Group 2, averaging eight hours and forty minutes, as compared to the latter's ten hours. Although the speed rates did not distinguish between application of the recognition and selection rules in the models, the catalogers' higher rates in applying the models would appear to be largely explained by the difficulties and frustrations they encountered in using the recognition rules, a point discussed earlier in the chapter.

In the case of both groups, however, the rules in the AA model were applied faster than those in the ALA model. A look at the scores shows that all but three of the participants applied these rules faster, ranging from eight minutes to almost two and a half hours. The fewer rules in the model, as well as the slightly less complex trees, an average of 7.3 nodes and 1.2 steps per rule as compared to 8.2 nodes and 2.3 steps in the ALA model (see Chapter V, Table 34, p. 236), would appear to account for the somewhat lower rates in application.

The comparability of the scores, for the models as well as the groups, presents a number of implications. The conclusions that can be drawn from the findings of this test and those of the first test, as well as the practical implications, will be discussed in the following chapter.

Table 48

Hours Taken Per Model

	AA Model	ALA Model	Total
Group 1:			
1	3.00	4.50	7.50
2	3.67	5.33	9.00
3	3.42	3.58	7.00
4	3.83	6.17	10.00
5	5.17	5.33	10.50
6	4.17	3.83	8.00
Average Number of Hours:	3.88	4.79	8.67
Average Minutes per Title Page:	10.42	13.22	
Group 2:			
1	5.75	5.25	11.00
2	4.17	4.83	9.00
3	6.33	7.17	13.50
4	2.92	3.08	6.00
5	4.17	3.83	8.00
6	5.00	7.15	12.15
Average Number of Hours:	4.72	5.22	9.94
Average Minutes per Title Page:	12.87	14.17	

4. Post-test questionnaire and interviews. — Although the post-test questionnaire had no bearing on the findings, the responses of the participants were considered to be of sufficient interest to report. As was pointed out earlier, except for Question 1, which concerned the catalogers, the remaining three questions were directed to all participants (see Appendix F, pp. 461-63). The ease with which the rules and the models were applied, and the extent to which they accommodated mechanical application elicited differing and, in some cases, contrasting responses from the two groups. For each question, there was a set of numbered responses, one of which the participants were asked to circle. Brief notation of the questions and a tally of the numbered responses are reported in Table 49.

In reference to the rules and the models, the majority of the participants in Group 1 indicated that there was no difference in the degree of ease with which either was applied. The catalogers, on the other hand, clearly distinguished between the rules and the models, with the majority considering the selection rules easier to apply than the recognition rules, and the AA model easier than the ALA model. Three of the catalogers pointed out the complexity of the recognition rules and their "inaccuracy", referring no doubt to the non-names they produced. Choice of the AA model was defended on the grounds that it was less tedious to apply, having fewer and shorter rules.

The responses of both groups with respect to the overall mechanicability of the rules were of particular interest. Rating the rules on a

Table 49

Responses to Questionnaire

Question	Group 1	Group 2
2. Application of recognition rules and selection rules		
1. RR easier than SR	2	
2. SR easier than RR		5
3. No difference	4	1
4. Do not remember		
3. Application of AA model and ALA model		
1. AA model easier than ALA model	1	4
2. ALA model easier than AA model		
3. No difference	5	2
4. Do not remember		
4. Mechanicability of models		
1. Extremely mechanical	4	1
2. Mechanical to large degree	2	4
3. Semi-mechanical		1
4. Mechanical to small degree		
5. Not mechanical at all		

five-point scale, the majority of the participants in Group 1 indicated that the rules were extremely mechanical, with comments to the effect that they required little or no thinking. In contrast, as the table shows, only one cataloger rated the rules extremely mechanical; four considered them mechanical to a large degree, and one judged them semi-mechanical. Participants who rated them in the latter two categories explained that they still used "judgment", especially in abstracting the strings. The participant, who rated them semi-mechanical, pointed out that they worked on the sample, but doubted that they would be successful in her own work.

Post-test interviews were arranged with the participants to discuss points in application. Except for the participant who distinguished the strings in the data on title page c, all the participants in Group 1 who made errors in application agreed that those errors were a result of oversight. The participants in Group 2 who failed to apply recognition rule 6 revealed their reluctancy to record non-author names. Other errors in the application of both sets of rules in the models were attributed to an oversight.

Summary. — The findings resulted from the test of mechanicability performed by two groups of participants fully supported the conclusion that both models were highly and almost equally mechanical. Application of the models to a problem set of title pages showed the scores of the group as well as the models to be comparable. With the exception of one error in the application of the recognition rules and selection rules, misapplication of both the recognition rules and selection rules was attributable to human error and oversight on the part of the participants. The results of their application are summarized at the ends of sections 1 to 4 in this chapter.

The conclusions that can be drawn from the findings of this test and those of the first test, as well as the practical implications, will be discussed in the following chapter.

CHAPTER VII

SUMMARY AND CONCLUSIONS

The purpose of this study was to determine whether the human intellectual process of selecting main entry can be simulated by automatic techniques. Simulating the process by automatic techniques referred specifically to the use of mechanistic devices and procedures, which would facilitate either machine or comparable automatic manipulation of cataloging data to produce appropriate selections of main entry. The machine technique used in the study to simulate the entry process was directed tree graphs. Rules of entry in both the AA and ALA cataloging codes were considered amenable to representation by this technique. However, differences in the content, number, design, and structure of the rules in these codes were seen to affect the complexity, operational capacity, and ability of the trees to make appropriate main entry selections. On this basis, investigation was made of the hypothesis that a corpus or group of rules for selection of entry in the AA code would be more amenable to simulation by the proposed technique than its counterpart corpus of rules in the ALA code.

The amenability of each corpus was determined in terms of four criterion variables: number, complexity, and redundancy of code rules, and the specific enumeration of the bibliographic conditions contained in those rules. The extent to which these variables affected the ability of the models to simulate the entry selection process was determined through the application of two tests of adequacy. The adequacy of the models to assign appropriate main entry selections was demonstrated in the first test, which involved the application of each model to a random sample of Library of Congress proof slips. The adequacy of the models to accomplish entry selection automatically, that is objectively and non-intuitively, was demonstrated in a second test involving an experiment, in which the simulation models were applied to a specially selected problem set of English language monographs. In the balance of this chapter, the findings of the study are summarized. The limitations and assumptions are noted, and some conclusions that can be drawn from the findings are outlined. Finally, some practical implications of the findings and conclusions are discussed, followed by suggestions for further research.

Summary of the Findings

1. With one exception, measurement of the criterion variables in the two sets of main entry rules did not support the contention of the hypothesis:

(1) Measurement of the variable of size showed the AA corpus with 19 rules to be a considerably smaller corpus to characterize than the counterpart ALA corpus with 36 or almost twice the number of rules to characterize.

(2) However, measurement of the variable of redundancy showed the AA corpus to be considerably more redundant than the ALA

corpus, with over 100 per cent repetition in the bibliographic conditions and specifications of entry in the rules in contrast to 63 per cent found in the rules in the ALA corpus.

(3) Measurement of the criterion variable of complexity showed the AA and ALA corpus to be almost equally complex with respect to the number of bibliographic conditions in their rules that required characterization.

(4) Measurement of the variable of specificity showed both sets of rules to be equally specific with respect to the number of their conditions that were exact, in the sense that they permitted observation in the title page data and subsequent characterization.

2. The results of the test of accuracy applied to the formal models constructed for the two sets of main entry rules did not support the prediction of the hypothesis:

(1) Application of the recognition rules in the AA and ALA models to the bibliographic data contained on Library of Congress proof slips showed both models to be highly successful and comparable in identifying the author names in the data - performing at the 93 and 95 per cent level respectively.

(2) Application of the experimental selection rules in the models to the author names identified by the recognition rules showed both models to be highly and equally successful in the assignment of correct main entries - performing at the 97 per cent level.

(3) Examination of the experimental selection rules for

their ability to assign the required types of main entry showed only slight differences in the models. Of the two types of entry, author and title, both models in general were found to have more difficulty assigning correct title main entries.

(4) Similarly, there were only slight differences in the two models with respect to the performance of their individual selection rules. Of the 9 AA and 11 ALA rules, all but four of the former and all but six of the latter assigned correct main entries. The remaining rules in the two models assigned incorrect as well as correct main entries. Furthermore, the pairs of similarly constructed rules in the models, of which there were nine, were found to be comparable in both performance and frequency of use.

(5) Correlation of the correct entries with their applicable rules of entry in each corpus showed that there were considerably more rules in the ALA corpus that had been successfully characterized, 22 or 61 per cent in contrast to 8 or 42 per cent in the AA corpus. However, when the rules in each corpus were correlated according to their comparability determined in the study, the evidence indicated that an almost equal number were successfully characterized by the two models.

(6) Examination of the incorrect entries assigned by the models showed that over half in both samples were due to unsuccessful identification of the author names, which resulted from application of the recognition rules. With suggested alterations in the recognition

rules, the author names on approximately half of these entries could be recognized successfully, and subsequently assigned correct main entries by the selection rules in the models.

(7) Approximately 21 per cent of the incorrect entries assigned by the AA model and 15 per cent of those assigned by the ALA model resulted from incomplete characterization of the entry rules, where certain environments associated with the patterns of authorship in these rules were not incorporated in the experimental selection rules. In the case of both models, a small and equal number of entry rules were characterized incompletely. By making the necessary adjustments in the selection rules of the two models, all of these entries could be assigned correct main entries.

(8) The remaining incorrect entries, which accounted for 18 per cent of those assigned by the AA model and 15 per cent of those assigned by the ALA model, were a result of deficient characterization of the entry rules, where bibliographic conditions in these rules were too conflicting or obscure for formal characterization. In the case of both models, a small and almost equal number of entry rules could not be characterized.

3. Examination of the tree diagrams showed that there were only slight differences in the two models with respect to the number of trees, nodes in the trees, and steps necessary to achieve entry using these trees.

4. Of the four criterion variables measured in the two sets of

main entry rules, two were found to be relevant factors of amenability.

(1) Examination of the characterizations in the two models showed that size or the number of rules in each corpus was not a relevant factor of amenability. While there was a considerable difference in the size of each corpus, this difference was not found or reflected in the characterizations developed in the study.

(2) Redundancy was seen to be a relevant factor of amenability. However, the quantitative measure posited for the variable in the study was found to be inadequate for determining its presence in the rules. Examination of the characterizations showed the AA corpus to be considerably less redundant than measurement of the variable indicated. Measurement of the variable showed the construct of one rule to be redundant with every other rule in this corpus. However, development of the characterization revealed this rule to be a "recycling" rule (a direction to begin the entry process again using the previous rules) rather than a redundant rule (an unnecessary repetition of the previous rules), thus reducing the redundancy in the AA corpus to a considerably smaller amount than was found in the ALA corpus.

(3) Examination of the characterizations in the models showed that complexity was a relevant factor of amenability. The number of bibliographic conditions in each corpus was reflected in the number of nodes in the characterizations representing these conditions. As in the case of the conditions in the entry rules, there were only slight differences in the number of nodes in the two characterizations.

(4) Similarly, specificity was found to be a relevant factor of amenability. Bibliographic conditions that were determined exact in each corpus were observed as specifications contained at the nodes in the characterizations. Both characterizations correspondingly reflected the findings measuring the variable in the rules. Measurement of the variable in the entry rules showed the two sets of rules to be equally amenable to simulation.

5. Finally, testing of the models with respect to their adequacy to accomplish selection of entry automatically showed both models to be almost equally mechanical:

(1) Application of the recognition rules in the AA and ALA models to a sample of title pages by two groups of participants showed the rules to be mechanical at the 90 and 95 per cent level respectively for Group 1, acting as automata in the study, and 82 and 77 per cent respectively for Group 2, comprised of experienced catalogers.

(2) Application of the experimental selection rules in the models to the bibliographic data identified by the recognition rules showed the AA and ALA models to be mechanical at the 91 and 95 per cent level respectively for Group 1, and at the 95 and 91 per cent level respectively for Group 2.

(3) With one exception, misapplication of the recognition rules by both groups of participants was a result of human error and oversight, where mis-reading, either the data or the rules, resulted in the omission or incorrect identification of the information on the

title page. The exception occurred in connection with a technicality in one rule, where failure to provide certain specifications resulted in a participant identifying the bibliographic information incorrectly.

(4) The lower scores in the application of both sets of recognition rules by the catalogers in Group 2 were found to result largely from a reluctance to identify non-author names, which could not be avoided in the case of certain rules.

(5) Similarly, with one exception, misapplication of the experimental selection rules by both groups of participants was a result of human errors made either in matching the recognitions to the construct of the rules, or through previous misapplication of the recognition rules. In the case of the exception, the technicality in the recognition rule noted in (3) above, accounted for the subsequent misapplication of a selection rule.

(6) With provision for the technicality and exclusion of the human errors noted, it was found that the scores for the mechanicability of both models would be increased to the 100 per cent level. However, the errors made by Group 2 were an important determinant of the effectiveness of the models in assisting catalogers to perform the entry selection process. Excluding the errors made in their application was irrelevant for determining their potential effectiveness for these people.

(7) The time taken to apply the two models to the sample showed both groups applied the rules in the AA model faster than those

in the ALA model, with a difference ranging from eight minutes to almost two and a half hours.

Limitations and Assumptions

The conclusions drawn from the findings summarized above presuppose a number of limitations and assumptions. Those that are most important to the study are restated below.

1. The trees and heuristics used in the development of the models describe, but do not constitute the operations of a computing mechanism performing the entry selection process.

2. Testing of the models for their accuracy and mechanicability was performed by humans. Findings applicable to human performance cannot be imputed to the performance of a computing mechanism.

3. The operational capacity of the models was limited to the identification and manipulation of the title and author statements on the title pages of English language books. There is no reason, however, to believe that it could not be extended to include the data on the title pages of non-English books.

4. The operational capacity of the models was further limited in application to title pages, where the title was the first or initial statement and the author was the second statement. While there is reason to believe that other physical patterns of these data accommodate

characterization, the findings based on this pattern do not allow generalization to other patterns.

5. Only a small number of main entry rules in the AA and ALA codes were characterized. Although the characterization developed for these rules was seen to provide the basis for the characterization of other main entry rules in these codes, the findings do not allow generalization to other rules.

Conclusions

Within the limitations, some conclusions can be drawn.

1. The study hypothesis which predicts that rules of entry in the AA code are more amenable to simulation by automatic techniques than counterpart rules in the ALA code was not supported by the data from this study. On the variables which were relevant for determining this amenability, the extremely small differences, which were found in the two codes, rejected the study hypothesis.

2. The contention that a model constructed for the main entry rules in the AA code will be more successful in assigning correct main entries than one constructed for counterpart rules in the ALA code finds no support in this study. The predicted differences in performance were not found, and the alternative hypotheses which were advanced further rejected the study hypothesis.

3. Similarly, the contention that a proposed model of AA rules of entry will be more mechanical than a proposed model of counterpart

ALA rules finds no support in this study.

4. There is evidence in this study to indicate that a formal model of entry selection is limited in its effectiveness to assist professionally trained and experienced catalogers to assign main entries mechanically. Knowledge and use of the rules of entry in the codes appears to hinder rather than assist mechanical application of a formal model.

5. In contrast, persons lacking library training and experience are seen to apply a formal model of entry selection with considerable accuracy and objectivity.

6. The study hypothesis which predicts that a model constructed for rules of entry in the AA code will be less complex than one constructed for counterpart rules in the ALA code receives very little support from the findings of this study. On all the measures applied to the models, differences were in favor of the AA model. However, these differences were extremely small.

7. At the same time, the overall economy achieved in the characterization of the AA rules was seen to account for this model's faster rate of application by both groups of participants in the study.

8. Data from this study indicate that successful characterization of the rules of entry in the AA and ALA codes is most closely associated with the number of bibliographic conditions in the rules that permit observation and identification in the data on the title page, and on the number of entries that are based on the physical display of these conditions.

9. To the extent that specificity and complexity were observed to affect the amount and kind of characterization in the models, these criterion variables are seen to be relevant factors of a code's amenability to automated simulation.

10. On the whole, correct assignment of the main entry was most difficult to achieve in the case of rules for multiple or shared authorship. In both codes, conflicting or obscure conditions of entry in these rules did not accommodate characterization.

11. Insofar as the choice of main entry in the AA and ALA codes is constructed on the concept of "responsibility in authorship", the philosophy is not seen to facilitate fully automated simulation (i.e. simulation without human intervention). The data from this study show that the knowledge and judgment of the cataloger are indispensable for complete interpretation of the concept. Although there is a stated attempt in the AA code to avoid dependence on the cataloger by underscoring the use of the title page as the source for the determination of the main entry, the data does not indicate that the cataloger is any less dispensable for the interpretation of the concept in this code than in the ALA code.

12. On the other hand, findings of this study indicate that for the majority of the works of personal authorship, assignment of the main entry can be accomplished on the basis of the information on the title page alone, with little or no dependence on the judgment of the cataloger.

13. The similarity of the model constructed for the AA code with that constructed for the ALA code indicates that the rules of entry for personal authorship in this code are little changed in content and application. The logical structure of this code and the purposive character of the principles on which it is based are seen to have resulted in rules of entry that are extrinsically, but not substantially, different from those in the ALA code.

Practical Implications

The findings and conclusions of this study have a number of practical implications, some of which are outlined below.

1. In libraries originating cataloging copy, cataloging supervisors should give more serious consideration to title page cataloging as a means of expediting the cataloging process. The number of works in the samples tested in the study, whose title page data accommodated correct assignment of the main entry, suggests that this form of cataloging might be used profitably for a large number of acquisitions received in libraries.

2. Further practical and economic benefits are to be derived in assigning this work to non-professional staff. In the random samples tested in the study, the majority of the entries involved the application of rules for single and joint or shared authorship. With basic instructions in the application of these rules alone, non-professional staff could provide at least the initial cataloging copy for a considerably large number of acquisitions. Although testing was limited, it

would appear that these people could also make productive use of the models in the study to provide cataloging copy for an even wider range of acquisitions.

3. In line with this approach, heads of cataloging departments should utilize the training and experience of professional staff members to supervise and review the work of the assisting staff, and to catalog more difficult materials. In the samples tested in the study, only a small number of entries were seen to require initial cataloging by the cataloger. The professional time and attention given to initiating copy for the remaining majority of entries would indicate that cataloging departments have failed to utilize catalogers efficiently and economically.

4. The implications of this study for library education are difficult to specify. However, if it is the case that the cataloging process or certain parts of it that are now performed by professional librarians can be automated or assumed by non-professional staff, then library schools should deemphasize teaching those skills that can be relegated to machines or these other people, and concentrate on preparing librarians to administer cataloging departments and the complex of staff and operations associated with the larger area of technical services, and to catalog those materials that require professional knowledge and skill.

5. Finally, if librarians are to consider a system in which automated cataloging is to play a part, it is essential that the intellectual

structure of the existing cataloging process be reexamined in light of the capabilities and operations performed by machines. The capabilities of the machine to record and provide access to a bibliographic item under numerous identification tags - author, title, date, publisher, etc. make the concept of main entry and the problems involved in determining that entry less urgent. Indeed, it is questionable whether rules for author and title entry will be required at all. However, given the tenability of the main entry as a mode for identifying bibliographic entries in single-access catalogs, such as union catalogs, booklists, etc., in bibliographic citations, and indeed, in every day communications, the data from this study point up the need to examine the relevancy of the concepts and procedures used to establish this entry.

Suggestions for Further Research

The specific area of this study opens up a vast field for further necessary research in the continuing examination of the role of automation in the cataloging process. Below are some ideas for replicating and extending the research of the present study.

In view of the exploratory nature of the study, the findings and conclusions are necessarily tentative. Other studies, either replicating the present study or along similar lines, should be done to validate, supplement, or contradict the findings and conclusions. Within the parameters of the present research, some needed refinements and improvements suggest themselves. The use of one random sample for testing the models is seen to permit a more precise evaluation of their performance. In turn, the random sample should preferably consist of title pages rather than the data recorded on Library of Congress proof slips. Although the stated assumptions in use of the latter were not seen to bias the findings, application of the models to various configurations of the data on title pages would provide more rigorous testing.

Conclusions as to the relevancy of the variables as factors of amenability must be considered very tentative. For instance, other more sophisticated measures are needed to determine the amount of redundancy in the rules, and the extent to which it reduces the amount of characterization in the models. Similarly, the extent to which complexity in the rules is exerted in the amount of characterization in the models needs more rigorous analysis. A base other than that of the structural descriptions might be used for measurement of the variables, perhaps direct use of the verbal statements in the rules. However, a difficulty in using the rules or deriving a base from them relates to the nature and construct of the rules themselves. As verbal directions, whose import is subject to a variety of interpretations, they provide a limited base for objective measurement.

Consideration should be given to using other and larger groups of persons to apply the models. As a substitute for freshman students, for instance, clerical workers, housewives, etc. could be considered. Other groups of professional personnel -- catalogers in regional processing centers, large public libraries, etc. could be substituted for those selected in the study.

Research on the amenability of the codes should be extended to include more rules of entry, both general and special. One might ask if the similarities in the models hold for other rules. The rules characterized in the study also need further analysis to determine the adequacy of the characterizations for other less general patterns of authorship associated with their use. In addition to implementing the models developed in the study, research should be directed toward designing and developing other models of the same or related cataloging processes.

Outside the parameters of this research, there is most obviously the need to study the technical and economic feasibility of the models developed in the study. Because of the number of operations and problems involved, this research is seen to require the coordinated efforts of librarians and computer scientists.

BIBLIOGRAPHY

American Library Association; Library of Congress, Library Association; and Canadian Library Association. *Anglo-American Cataloging Rules*. North American Text. Chicago: American Library Association, 1967.

American Library Association. Division of Cataloging and Classification. *A.L.A. Cataloging Rules for Author and Title Entries*. 2d ed. Edited by Clara Beetle. Chicago: American Library Association, 1949.

Avram, Henrietta D.; Guiles, Kay D., and Meade, Guthrie T. "Fields of Information on Library of Congress Catalog Cards: Analysis of a Random Sample 1950-1964." *The Library Quarterly*, XXXVII (April, 1967), 188.

Baker, Norman R., and Nance, Richard E. "The Use of Simulation in Studying Information Storage and Retrieval Systems." *American Documentation*, XIX (October, 1968), 363-70.

Bar-Hillel, Jehoshua. *Language and Information*. Reading, Mass.: Addison-Wesley Publishing Company, 1964.

Baxendale, Phyllis. "Machine-Made Index for Technical Literature--An Experiment." *IBM Journal of Research and Development*, II (October, 1958), 354-61.

Caras, Guy J. "Computer Simulation of a Small Information System." *American Documentation*, XIX (April, 1968), 120-22.

Chapman, R. L. "The Case for Information System Simulation." *Information System Sciences*. Washington, D.C.: Spartan Books, 1964.

Cuadros, C. A., ed. *Annual Review of Information Science and Technology*. Chicago: Encyclopedia Britannica, 1966-70.

Current Research and Development in Scientific Documentation. Washington, D.C.: National Science Foundation. Office of Scientific Information, 1957-69.

Field, F. Bernice. "Anglo-American Cataloging Rules Correlated with A.L.A. Cataloging Rules." *New Rules for an Old Game*. Edited by Thelma E. Allen and Daryl L. Dickman. Vancouver: University of British Columbia, 1967.

Field, F. Bernice. "The Rules for Entry and Heading." New Rules for an Old Game. Edited by Thelma E. Allen and Daryl L. Dickman. Vancouver: University of British Columbia, 1967.

Bull, C. D. "How Will Electronic Information Systems Affect Cataloging Rules?" Library Resources & Technical Services, V (Spring, 1961), 135-39.

__________. "[Working Paper] No. 17, The Impact of Electronics Upon Cataloguing Rules." The Report of the International Conference on Cataloguing Principles, Paris, 9th-18th October, 1961. London: International Federation of Library Associations, 1963.

Hays, David G. Introduction to Computational Linguistics. New York: American Elsevier Publishing Company, 1967.

Hunt, Earl B.; Marin, Janet; and Stone, Phillip J. Experiments in Induction. New York: Academic Press, 1966.

Jakobson, Roman, and Halle, Morris. Fundamentals of Language. The Hague: Mouton & Co., 1956.

Language and Machines: Computers in Translation and Linguistics. Washington, D.C.: National Academy of Sciences, 1966.

Lubetzsky, Seymour. "Catalog Code Review, 1964." Library Journal, LXXXIX (December 15, 1964), 4863-65.

Melcher, Daniel. "Cataloging, Processing, and Automation." American Libraries, II (July-August, 1971), 703.

Montgomery, Christine, and Swanson, Don R. "Machinelike Indexing by People." American Documentation, XIII (October, 1962), 359-66.

Proceedings of the Second Conference on Machine-Readable Catalog Copy, November 22, 1965. Washington, D.C.: Library of Congress, 1965.

Rapoport, Anatol. "Mathematical Models of Social Interaction." Handbook of Mathematical Psychology. Edited by R. Duncan Luce, Robert R. Bush, and Eugene Gallanter. Vol. II. New York: John Wiley and Sons, Inc., 1963.

Letter from John C. Rather, Specialist in Technical Processes Research, Library of Congress, October 8, 1969.

Ruhl, M. J. "Chemical Documents and Their Titles: Human Concept Indexing Versus KWIC-Machine Indexing." American Documentation, XV (April, 1964), 136-41.

Salton, Gerard. Automatic Information Organization and Retrieval. New York: McGraw-Hill Book Co., 1968.

__________. "The Manipulation of Trees in Information Retrieval." Information Storage and Retrieval System. Report ISR-1, Section IV. Cambridge, Mass.: Computation Laboratory, Harvard University, 1961.

Slocum, Robert B., and Hacker, Lois. Sample Cataloging Forms. 2d ed. Metuchen, New Jersey: Scarecrow Press, 1968.

Stevens, Mary Elizabeth. Automatic Indexing: A State-of-the-Art Report. Washington, D.C.: Government Printing Office, 1965.

Swanson, Don R. "Library Goals and the Role of Automation." Special Libraries, LIII (October, 1962), 470.

__________. "Searching Natural Language Text by Computer." Science, CXXXII (October 21, 1960), 1099-1104.

Tait, James A. Authors and Titles. Hamden, Conn.: Archon Books, 1969.

U. S. Library of Congress. Descriptive Cataloging Division. Rules for Descriptive Cataloging. Washington, D.C.: Government Printing Office, 1949.

U. S. Library of Congress. Processing Department. The Cataloging-In-Source Experiment. Washington, D.C.: Library of Congress, 1960.

__________. Cataloging Service, Bulletin No. 65 (August, 1964), 3.

__________. Cataloging Service, Bulletin No. 79 (January, 1967), 1.

__________. Cataloging Service, Bulletin No. 80 (April, 1967), 1.

__________. Cataloging Service, Bulletin No. 88 (January, 1970), 1.

Vickery, B. C. On Retrieval System Theory. London: Butterworths, 1965.

Wasserman, Paul; Daniel, Evelyn, and Logan, Kay, eds. LIST 1971: Library and Information Science Today: An International Registry of Research and Innovation. New York: International, Inc, 1971.

APPENDIX A

RULES OF ENTRY SELECTED FOR CHARACTERIZATION

RULES OF ENTRY CHARACTERIZED IN THE STUDY

The American Library Association, and Others. Anglo-American Cataloging Rules. North American Text. Chicago: American Library Association, 1967.

GENERAL RULES

1. Works of single authorship

A. Enter a work, a collection of works, or selections from works by one author under the person or corporate body that is the author, whether named in the work or not.

2. Works of unknown or uncertain authorship, or by unnamed groups

A. Enter under title a work that is of unknown or uncertain authorship, or that is by a group that lacks a name. Make added entries under as many as three persons to whom authorship has been attributed, either in editions of the work or in reference sources. Always make an added entry under a person to whom authorship is attributed in the edition being cataloged.

C. If the only clue to authorship is in the appearance on the title page of initials, some other alphabetical device, a characterizing word or phrase preceded by the indefinite article, or a phrase naming another work that the author wrote, make an added entry under this clue.

3. Works of shared authorship

Preliminary note. The rules below apply to 1) works produced by the joint collaboration of two or more authors, 2) works for which different authors have prepared separate contributions (e.g. composite works, symposia, series of addresses, lectures, etc., written specifically for a particular occasion or for the publication in hand) providing the authors are not acting as members of a corporate body (see 17), and 3) works consisting of an exchange between different persons (e.g. correspondence, debates). They apply equally to cases of shared responsibility among compilers, editors, translators, adapters, etc., when the applicable rule requires main entry under such a person.

For special types of collaboration see 8A, 13, and 16. When a publication results from the work of persons under the direction of a supervisory editor, apply rule 4. For collections of previously existing works by different authors apply rule 5.

A. Principal author indicated. Enter a work of shared authorship under the person or corporate body, if any, to whom principal responsibility is attributed, e.g. by wording or typography. Make added entries under the other authors involved if there are not more than two. Always make an added entry under an author, other than the principal author, whose name appears first on the title page.

B. Principal author not indicated

1. If no one is represented as principal author and if there are not more than three authors, enter under the one that is named first and make added entries under the others.

2. If no one is represented as principal author and there are more than three authors, enter the work under its title unless the work is produced under the direction of an editor named on the title page. In this case apply rule 4. Make an added entry under the author named first on the title page.

4. **Works produced under editorial direction**
Preliminary note. Many works of multiple authorship are produced under the direction of an editor, who may or may not be named. Some of these works are serials, others are not. Some involve questions of corporate authorship, others do not. This rule applies only to such of these works as are not serials and do not involve questions of corporate authorship, other than the responsibility of the publisher. For serials see 6. For works involving questions of corporate authorship see 17.

A. Enter a work produced under editorial direction under its editor providing: 1) he is named on the title page of the work, 2) the publisher is not named in the title, and 3) the editor appears to be primarily responsible for the existence of the work. In all other cases enter under the title and make an added entry under the editor if he is named on the title page. If authors are named on the title page, make an added entry under the one named first.

5. Collections

A. With collective title. If a collection of independent works by different authors, not written specifically for the same occasion or for the publication in hand, or of extracts from such works, has a collective title, enters it under the person who compiled it when he is named on the title page as compiler or editor, or is known to have been so named in another edition of the collection; otherwise enter it under [illegible] title. Add the designation comp. to the heading. Use this designation also if the compiler translated the collection. Make an added entry under the compiler, editor, or title that is not used as main entry. Make author-title added entries for the works included if there are not more than three. If there are more than three works but not more than three authors, make added entries under each for whom an author-title added entry would be inappropriate. Make an added entry under a corporate sponsor that is named on the title page but not under a body created expressly to oversee the preparation of the publication.

B. Without collective title. If a collection of previously existing works by different authors, or of extracts from such works, lacks a collective title, enter it under the heading appropriate to the work listed first on the title page, or, if there is no collective title page, under the heading appropriate to the first work in the collection. Make an added entry under the compiler or editor if he is named on the title page as compiler or editor. Make added entries under the works or authors included if appropriate under the criteria given in A above. For the treatment of anonymous composite manuscripts and manuscript groups without title see 103.

WORKS WITH AUTHORSHIP OF MIXED CHARACTER

7. Adapter or original author

A. Enter an adaptation or other rewriting of a work in a different literary style (e.g. a paraphrase, epitome, version for children) or in a different literary form (e.g. a dramatization, novelization, versification) under the person who did the adapting or rewriting, if known; otherwise under its title. Make the appropriate added entry for the original work. In the case of paraphrases of Biblical texts, however, the added entry is omitted; a form subject heading is provided instead. For the treatment of translations see 15. For condensations and abrdiged editions in which the text has been reduced by the omission of passages or sections see 14A.

8. Artist or author of text

   A. Collaborative work. Enter a work that is or appears to be a work of collaboration between an artist and the author of the text under the one who is named first on the title page unless the other's name is given greater prominence by wording or typography. In case of doubt, prefer entry under the author of the text. Make an added entry under the one not given main entry.

   B. Illustrated work. Enter a work for which an artist has provided illustrations under the author of the text. Make an added entry under the illustrator only if the illustrations are a noteworthy feature of the work.

9. Biographer/critic or author

   A. Enter a publication consisting of a work or group of works of an author, accompanied by or interwoven with biographical or critical material by another person, under the latter if he is represented as author. Make an author-title added entry under the author whose work is published with the biographical or critical material. In case of doubt enter under the person whose name appears first.

   B. If the one who is responsible for the biographical or critical material is represented as editor, enter under the author of the work and made an added entry under the editor.

14. Reviser or original author

   A. In general, enter an edition that has been revised, enlarged, abridged, condensed, etc., by another person, under the original author, with an added entry under the other person.

   B. If the new edition clearly indicates that the work is no longer that of the original author, however, enter it under the new author. Make an author-title added entry under, or an explanatory reference from, the heading for the earlier work.

15. Translator or author

A. Enter a translation of a work under the author of the original work. Make an added entry under the translator according to the provisions of 33E. If the translation involves adaptation or is described as a "free" translation, however, treat it as an adaptation (see rule 7).

16. Writer or nominal author

A. Enter a work written by one person for another person in whose name it is presented under the latter person, with an added entry under the person who did the writing if his name appears openly in the work. See also 13C.

RELATED WORKS

19. Related works

B. Other related works. Enter any other related work under its own author and/or title according to the general rules. Make an added entry under the author and title or under the title of the work to which it is related, as appropriate.

American Library Association, Division of Cataloging and Classification. A.L.A. Cataloging Rules for Author and Title Entries. 2d ed. Edited by Clara Beetle. Chicago: American Library Association, 1949.

2. Individual author, Works by.

Enter a work under the name of its author, when known, whether or not his name appears in the publication. (Cf. 32.) Enter revisions and other modifications under the original author whenever the work remains substantially his, especially if the book purports to be an edition of the original work. (Cf. 20-22.)

3. Joint authors, Works by.

A. Enter under the first author mentioned on the title page a work produced jointly by two or more authors in which the contribution of each is not a separate and distinct part of the whole. Make added entry with designation JOINT AUTHOR for any author after the first whose name is included in transcribing the title.

B. In a work of joint authorship in which the chief responsibility rests with one author, but the title page reads "with the collaboration of . . ." or words to that effect, the making of added entries for the collaborators will depend on (1) the nature of the work; (2) the number of collaborators and the importance of their contribution.

E. Enter a narrative told by one person (reporter, "ghost writer," etc.) who prepares the matter for publication, under the narrator. Make added entry for the writer.

Distinguish from works in which the narrator is the actual author those in which the writer has merely used the narrator form for an original work (e.g., autobiography of a fictitious person, or fictitious memoirs of a real person).

4. Composite works.
Enter composite works (Cf. GLOSSARY.) under the

author chiefly responsible for the work, under the author first mentioned on the title page, or under title, according to the specific rules A-B.

A. Collaborators, Works by.

(1) Enter a work produced by the collaboration of two or more authors, in which the contribution of each forms a distinct part or section of a planned whole, under the author chiefly responsible for it. Whenever their contributions warrant, make added entry for each of the other contributors without the designation JOINT AUTHOR. If a corporate body is considered to be chiefly responsible for the work, make an added entry for at least the first personal author mentioned on the title page.

(2) If origin, chief interest, or responsibility is not clearly identified with or attributable to any one of the contributors, enter under the first-named author if there are not more than three and the title of the whole work is applicable to each of the contributions, with added entry for the other authors. Otherwise, enter under title, with added entry for the first author mentioned and for as many others as the individual case warrants.

B. Debates, conversations, interviews, table-talk.

(1) Enter a debate, dialogue, conversation, under the participant whose name appears first on the title page. Make added entry for each name after the first.

5. Collections and serials.
Specification.—This group of rules applies to material in which the responsibility for the intellectual content is so divided that no one of the contributors can be considered the chief author. The choice then is between entry under the agency responsible for editing, assembling or issuing the material, as a government agency, society, institution (Cf. 71.), a personal editor, a publisher, etc., or entry under title. Preference is given to entry under title when

it is distinctive or when it is the least variable element and it is, of course, the only possible entry when no editorial or publishing responsibility is apparent.

A. Collections.

(1) With collective title. Enter under the compiler or editor, individual or corporate, a collection of independent works by various authors, artists, composers, etc., issued with a collective title, except as noted in (a) to (c) below.

(a) Enter such collections under title if the work of the editor or editing body seems to be but slight and their names do not appear prominently in the publication or if there are frequent changes of editor. Make added entry under editor.

(c) Enter festschriften and similar collections published by a society or an institution in honor of a person, or to celebrate an anniversary, under the society or institution. When not published by a society or an institution, enter under editor, if the editor's name appears prominently in the publication; otherwise enter under title. In any case, make added entry under editor or distinctive title when not main entry, and make subject entry under the person in whose honor the collection is published.

(2) Without collective title. Enter two or more writings by different authors, published together but having no collective title, under the one named first on the title page, or lacking such indication, under the heading appropriate to the first work in the collection. Make added entries for other works mentioned on the title page and for the editor.

(B) Encyclopedias and dictionaries. Enter encyclopedias and dictionaries under title unless decidedly better known by the name of their editors. In either case make added entry under the form not chosen for main entry, and for the publisher if the work is likely to be referred to by his name.

(C) Periodicals and newspapers

(4) Enter a collection of extracts by various authors from a single periodical or newspaper under the name of the periodical or newspaper. Make added entries under the title of the collection and the name of the collector if given.

For entry of a collection of extracts by various authors from different periodicals see 5A.

The work of a single author republished from a periodical (entered under the name of the author if known, or under title if anonymous) has no added entry for the periodical, but an added entry is usually made under the newspaper for the work of a single author republished from it, especially if an author is an editor or official correspondent, or if the article is anonymous.

(E) Directories

(2) Enter a nonserial directory under the name of the compiler if mentioned on the title page. If no compiler is mentioned, enter under title.

WORKS OF SPECIAL TYPE

6. Correspondence and letters.

A. Enter correspondence between two or more persons under the first one mentioned on the title page, with added entry for the other party or parties to the correspondence, and for the editor or compiler.

B(1) If the letters are all addressed to the same person, make added entry for the addressee. Extend this treatment to include two writers or two recipients. When more than two writers are involved apply rule C, p. 17.

(2) When the collected letters of a single writer are accompanied by biographical matter and the whole is entitled "Life and letters," or words to that effect, choice of entry under the writer of the letters or under the author of the biography is governed by the relative extent and importance or interest of the letters as compared with the biography. In case of doubt,

prefer entry under the author of the biography. If entry is made under the writer of the letters, make added entry under the author of the biography; if under the author of the biography, make subject entry under the writer of the letters. If the latter is a very voluminous writer, both a subject and an added entry may be made.

C. Enter collected letters by various writers under the compiler or editor. (Cf. 5A.) If the letters are all addressed to the same person, make added entry for the addressee. Extend this treatment to include two recipeints. (Cf. B(1) above.)

14. Radio scripts.

A. Enter a radio script or series of scripts under the author, when known. If all of the series were written for a named program or for a single broadcasting company, make added entry under the name of the program followed by RADIO PROGRAM in parentheses, or under the name of the broadcasting company.

16. Ships' logs.

Enter the official log of a ship under the name of the ship. Make added entry for the commanding officer if his name appears prominently on the title page.

19. Works of art.

D. When an artist and an author collaborate in the production of a work, enter under the one whose contribution forms the predominant feature of the work, and make added entry for the other. In case of doubt, prefer entry under the author of the text.

E. When the contribution of the artist is limited to the illustration of a work independently written by another person, make added entry for the illustrator only if the illustrations are a noteworthy feature of the work.

But if the illustrations are assembled and published without the text, enter under the illustrator with subject entry under the name of the author of the text with subheading ILLUSTRATIONS.

WORKS RELATED TO PREVIOUS PUBLICATIONS

20. Revisions.

Enter a revision under the name of the original author with added entry for the reviser. (Cf. 2.)

The rule applies as long as the work remains substantially that of the original author, especially if it purports to be an edition of his work. If, however, the revision is substantially a new work, enter it under the reviser with added entry for the author and title of the original work.

21. Translations.

A. Enter a translation under the author of the original work, with added entry for the translator. (Cf. 32G, 33A.)

22. Abridgments, adaptations, excerpts.

A. Enter an abridgment, epitome, or outline under the same heading as the original work with added entry for the abridger, epitomizer, etc.

B. Enter an adaptation, a paraphrase, or a free translation, as a general rule, under the same heading as the original work with added entry for the adapter, paraphraser, etc. (Cf. 33D.)

However, if the adaptation or paraphrase has been so freely made as to bear slight kinship with the original work as in the case of many adaptations for the use of juvenile readers, or if it has become a classic in its own right, prefer entry under the adapter or paraphraser, with added entry under the author and title of the original work. In doubtful cases, prefer entry under heading used for the original work.

When entry is under adapter or paraphraser and the paraphrase is not limited to a single work, or if the work paraphrased is by a voluminous author under whose name there is a systematic arrangement of material in the catalog, make added entry under the author's name with subheading PARAPHRASES, TALES, ETC.

C. Enter excerpts (selections) from a single author under that author, with added entry under compiler.

For entry of excerpts, etc., from several authors see 5A.

23. Dramatizations and novelizations.

A. Enter a dramatization based on a novel, legend, poem or other literary form under the playwright, with added entry under the author and title of the work upon which the dramatization is based.

For musical dramas, see the rule for librettos, 12A(2).

B. Enter a novelization based on a play or other literary form under the novelist, with added entry under the author and title of the work upon which the novelization is based.

24. Parodies and imitations.

Enter a parody or an imitation under its own author with subject entry for the author and title of the work parodied or imitated.

For parodies not limited to a single work or for a parody of a single work by a voluminous author under whose name in the catalog there is a systematic arrangement of material, make subject heading under the name of the author parodied with subheading PARODIES, TRAVESTIES, ETC.

25. Continuations (Supplements).

A. Enter a continuation which is in the form of an independent work with author and title differing from that of the original work under its own author. Make added entry under the name of the original author followed by a brief title of his work, or, if preferred, a reference may take the place of an added entry.

C. Enter a single monograph issued as a supplement to a periodical under the author, with added entry for the periodical, followed by the word Supplement.

26. Sequels.

Catalog a sequel as an independent work. If the sequel is by a different author, make an added entry under the author of the work which it continues, followed by a brief title thereof.

28. Concordances.

Enter a concordance under its compiler with subject entry under the heading appropriate to the work concordanced.

32. Anonymous works (General).
Specification.---A strictly anonymous work is one in which the author's name does not appear anywhere in the book. The term is here extended to include (1) works in which the author's name appears in the book only in a concealed manner, so that the authorship is not obvious; (2) works whose authorship is indicated by a descriptive or generic word or phrase preceded by an article, e.g., "by a lover of justice," "by a physician," "by a bishop of the Church of England"; (3) works in which, instead of his name, the author gives the title of another of his books, e.g., "by the author of . . ."; (4) those in which the author uses initials, asterisks, or other symbols instead of his name.

Works in which an author uses as a name a specific or phrase with or without a definite article are treated as pseudonymous. (Cf. 30.)

See also Anonymous classics (33-35).

General rule.

Enter works published anonymously under author when known. Make added entry under title and added entry or reference under any phrase used instead of the author's name, or if preferred, instead of the first added entry make reference:

If the author is not known, enter under title, making added entry under the phrase expressing authorship.

If the phrase expressing authorship is devoid of meaning when removed from the context of the title, no reference or added entry is necessary.

B. "By the author of . . ." When an anonymous work of undiscovered authorship bears on its title page such a phrase as "by the author of," enter it under its title with an added entry under the title quoted, followed by the words AUTHOR OF. If different titles are thus used by the author at various times, make the added entry under the title most frequently referred to, or, in case of doubt, under one of his best known or earlier works. Refer from the title of each work to the title thus chosen, using the following form: "For other works by the same author see . . ."

CORRELATION OF AA AND ALA RULES OF ENTRY SELECTED FOR CHARACTERIZATION

Anglo-American Cataloging Rules: 19	ALA Cataloging Rules: 36
1A	2
	16
	22C
2A	32
2C	32
	32B
3A	3A
	3B
	4A(1)
3B1	4A(2)
	4B(1)
	6A
	6B(1)
3B2	4A(2)
4A	5A(1)
	5A(1c)
	5B
	5E(2)
5A	5A(1)
	5A(1a)
	6C
5B	5A(2)
7A	22B
	23A
	23B
	24
8A	19D
8B	19E

Anglo-American Cataloging Rules. 19	ALA Cataloging Rules: 36
9A	6B(2)
9B	6B(2)
14A	20 22A
14B	20
15A	21A
16A	3E
19B	5C(4) 14A 25A 25C 26 28

APPENDIX B

EXPERIMENTAL RULES

AA Recognition Rules

RR 1(a) Recognize BY in any substring

Recognition YES ☐ NO ☐

If YES, copy down recognition for the rule, set off with appropriate boundary markers in the appropriate string

____

____

Go to RR 1(b)

If NO, go to RR 1(b)

RR 1(b) Recognize BY as implied at the head of any substring in which there is no physical expression of the marker in it

Recognition YES ☐ NO ☐

If YES, copy down recognition for the rule with BY in brackets, [ ], set off with appropriate boundary markers in the appropriate string

____

____

Go to RR 1(c)

If NO, go to RR 1(c)

RR 1(c) Recognize

$$\left\{\begin{array}{l}\text{drawings}\\ \text{edited}\\ \text{foreword}\\ \text{illustrat(ed)(ions)}\\ \text{introduc(ed)(tion)}\end{array}\right\} \Big/ \_\_\text{BY}$$

in any substring

Recognition YES ☐ NO ☐

If $\boxed{\text{YES}}$, copy down recognition for the rule and carry recognition for the previous rule forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 2(a)

If $\boxed{\text{NO}}$, go to RR 2(a)

RR 2(a) Recognize $(\text{A})\_\_\_\overset{\&}{\text{AND}}\_\_\_\text{'S,S'}$ when

$$(\text{A})\_\_\_\overset{\&}{\text{AND}}\_\_\_\text{'S,S'} \;\Big/\; \_\_\_\ldots\text{bY}\ [\text{bY}]\ \ldots$$

Recognition $\boxed{\text{YES}}$ ☐ $\boxed{\text{NO}}$ ☐

If $\boxed{\text{YES}}$, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 2(b)

If $\boxed{\text{NO}}$, go to RR 2(b)

RR 2(b) Recognize $(\text{A})\_\_\_\text{'S,S'}$ when

$$(\text{A})\_\_\_\text{'S,S'} \;\Big/\; \_\_\_\ldots \left\{\begin{array}{l}\text{drawings}\\ \text{foreword}\\ \text{illustrat(ed)(ions)}\\ \text{introduc(ed)(tion)}\\ \text{to...edited}\end{array}\right\} \ldots\text{BY}\ [\text{bY}]\ldots$$

Recognition $\boxed{\text{YES}}$ ☐ $\boxed{\text{NO}}$ ☐

If $\boxed{\text{YES}}$, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 2(c)

If NO, go to RR 2(c)

RR 2(c) Recognize (A)___'S,S' when it occurs without bY [BY] in the substring

Recognition YES ☐ NO ☐

If YES, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 3

If NO, go to RR 3

RR 3 Recognize OF when

OF / Life and {
articles
conversations
correspondence
dialogues
diary, diaries
essays
journal, journals
letters
poems
screenplays
works
writings
} ___...

Recognition YES ☐ NO ☐

If YES, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 4

If $\boxed{\text{NO}}$, go to RR 4

RR 4 Recognize OF, FROM, BETWEEN, WITH when

$$\left\{\begin{matrix}\text{OF}\\ \text{FROM}\\ \text{BETWEEN}\\ \text{WITH}\end{matrix}\right\} \Big/ \left\{\begin{matrix}\text{articles}\\ \text{conversations}\\ \text{correspondence}\\ \text{dialogues}\\ \text{diary, diaries}\\ \text{essays}\\ \text{journal, journals}\\ \text{letters}\\ \text{poems}\\ \text{screen plays}\\ \text{works}\\ \text{writings}\end{matrix}\right) \underline{\quad}\ldots$$

Recognition YES ☐ NO ☐

If $\boxed{\text{YES}}$, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 5

If $\boxed{\text{NO}}$, go to RR 5

RR 5 Recognize OF after BY in the substring when

$$\text{OF} \Big/ \text{BY}\ [\text{BY}]\ldots\text{WITH}\ldots\underline{\quad}$$

Recognition YES ☐ NO ☐

If $\boxed{\text{YES}}$, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 6(a)

If NO, go to RR 6(a)

RR 6(a) Recognize AND, as implied, at the end of each string except the final string in the substring when:

(a) BY [BY] is at the head of the substring and there is no other physical expression of the marker BY or the markers AND, OF, or WITH in the substring

(b) BY [BY] is at the head of the substring and EDITED BY is at the end of the substring and there is no other physical expression of the marker BY or the markers AND, OF, or WITH in the substring

Recognition	YES	NO

If YES, copy down recognition for the rule with AND in brackets, , and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 6(b)

If NO, go to RR 6(b)

RR 6(b) Recognize AND, as implied in the form of a comma (,) or slash (/) in the substring, when:

(a) BY [BY] is at the head of the substring and there is no other physical expression of the marker BY or the markers OF or WITH in the substring

(b) BY [BY] is at the head of the substring and EDITED BY is at the end of the substring and there is no other physical expression of the marker BY or the markers OF or WITH in the substring

Recognition YES ☐ NO ☐

If YES, copy down recognition for the rule with AND in brackets, [ ], and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 6(c)

If NO, go to RR 6(c)

RR 6(c) Recognize AND, &, after BY in the substring, except when:

(a) $\left\{\begin{matrix}\text{AND}\\ \&\end{matrix}\right\}$ / BY [BY] ...WITH...___

(b) $\left\{\begin{matrix}\text{AND}\\ \&\\ [\text{AND}]\end{matrix}\right\}$ / BY [BY] ...$\left\{\begin{matrix}\text{edited\_\_\_\_\_}\\ \text{\_\_\_\_\_edited}\end{matrix}\right\}$

Recognition YES ☐ NO ☐

If YES, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 7

If NO, go to RR 7

RR 7 Recognize WITH after BY in the substring except when:

(a) WITH / BY [BY] ...___...BY [BY]

(b) WITH / BY [BY] ...___...OF

	YES	NO
Recognition		

If $\boxed{\text{YES}}$, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

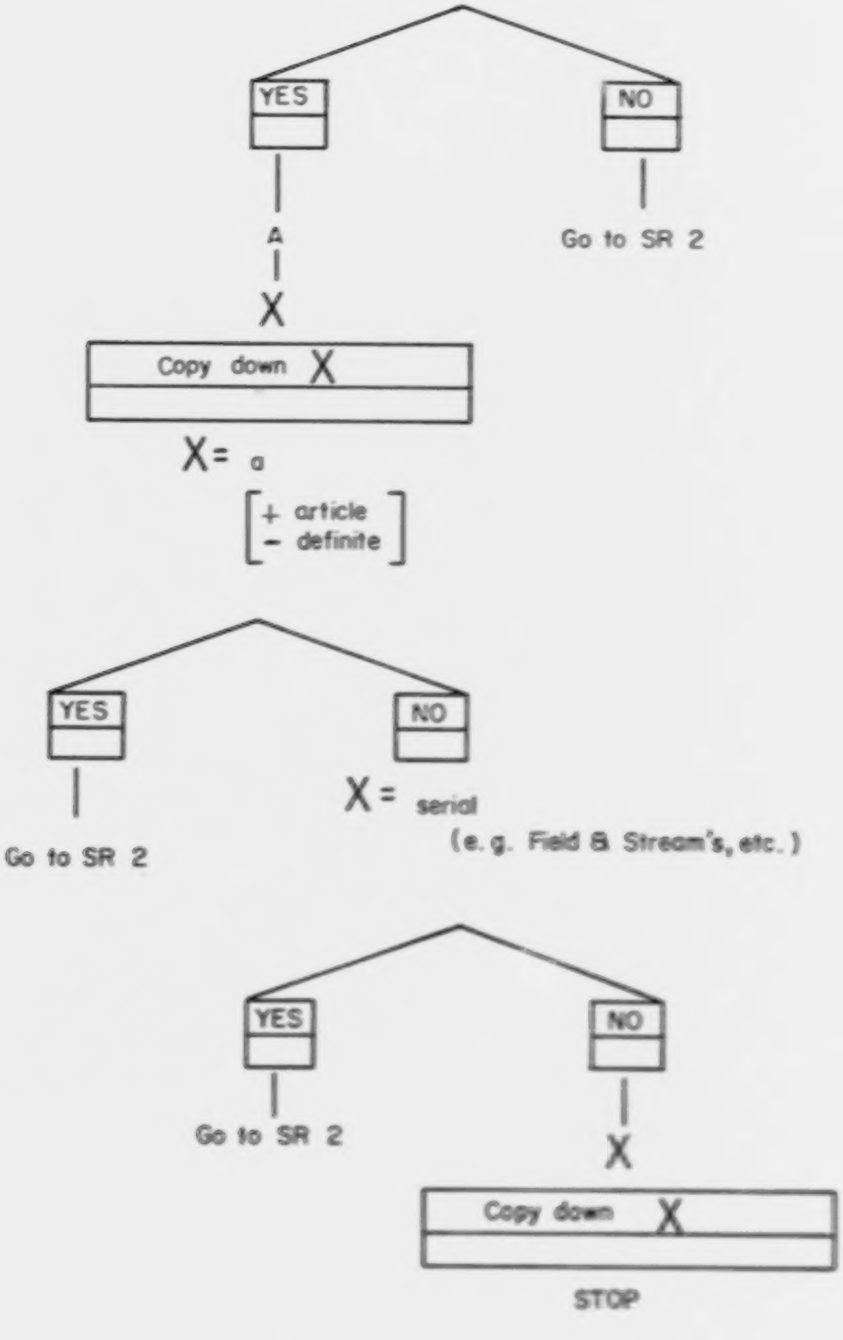
Apply AA Selection Rules to the recognition

If $\boxed{\text{NO}}$, apply AA Selection Rules to the last recognition copied down under the rules above or to the "non" or no recognition, as the case may be.

AASR 1
Recognition:

X $\overset{\&}{\text{AND}}$ **Y** 's, s' **Z** BY **WZ**

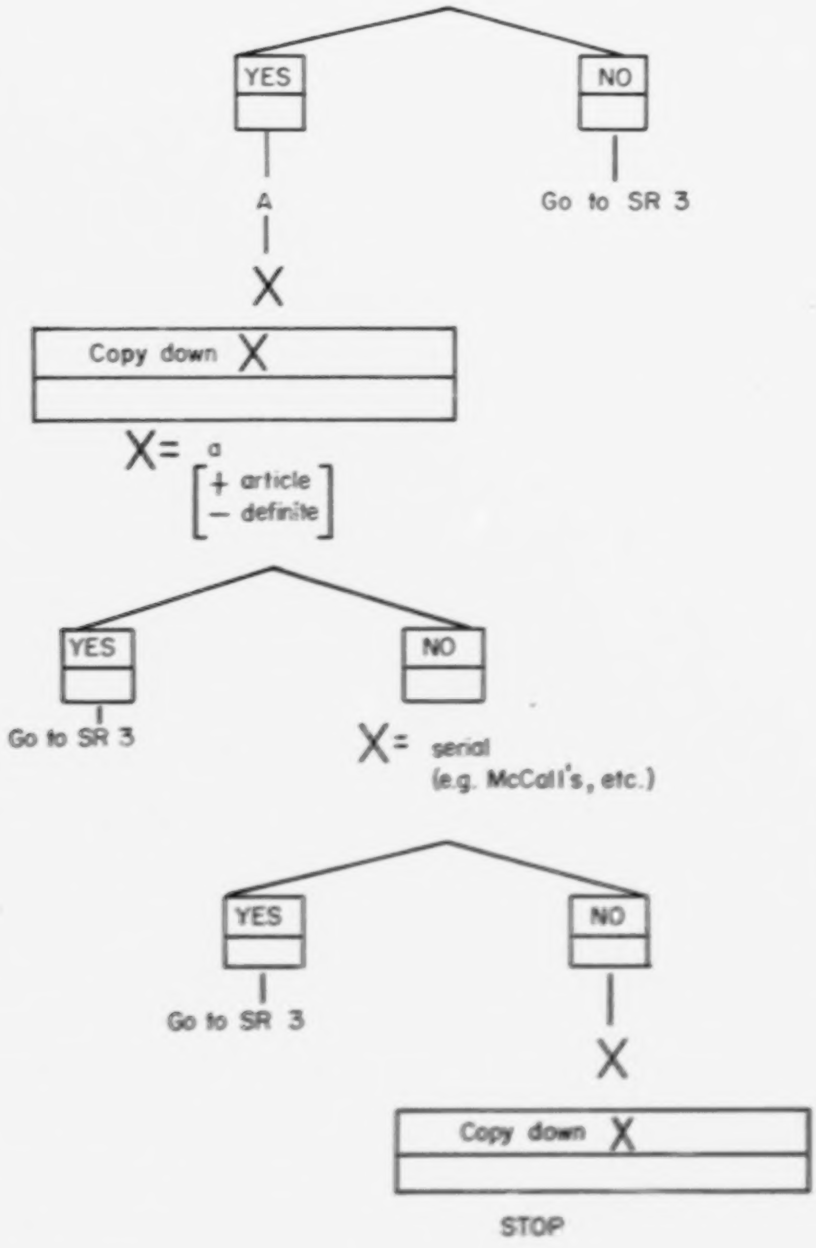
Where Z = anything or nothing


YES
NO
A
X
Go to SR 2
Copy down X
X = a
[+ article
− definite]
YES
NO
Go to SR 2
X = serial
(e.g. Field & Stream's, etc.)
YES
NO
Go to SR 2
X
Copy down X
STOP

AASR 2
Recognition:

$X_{\text{'s, s'}}\ Z$

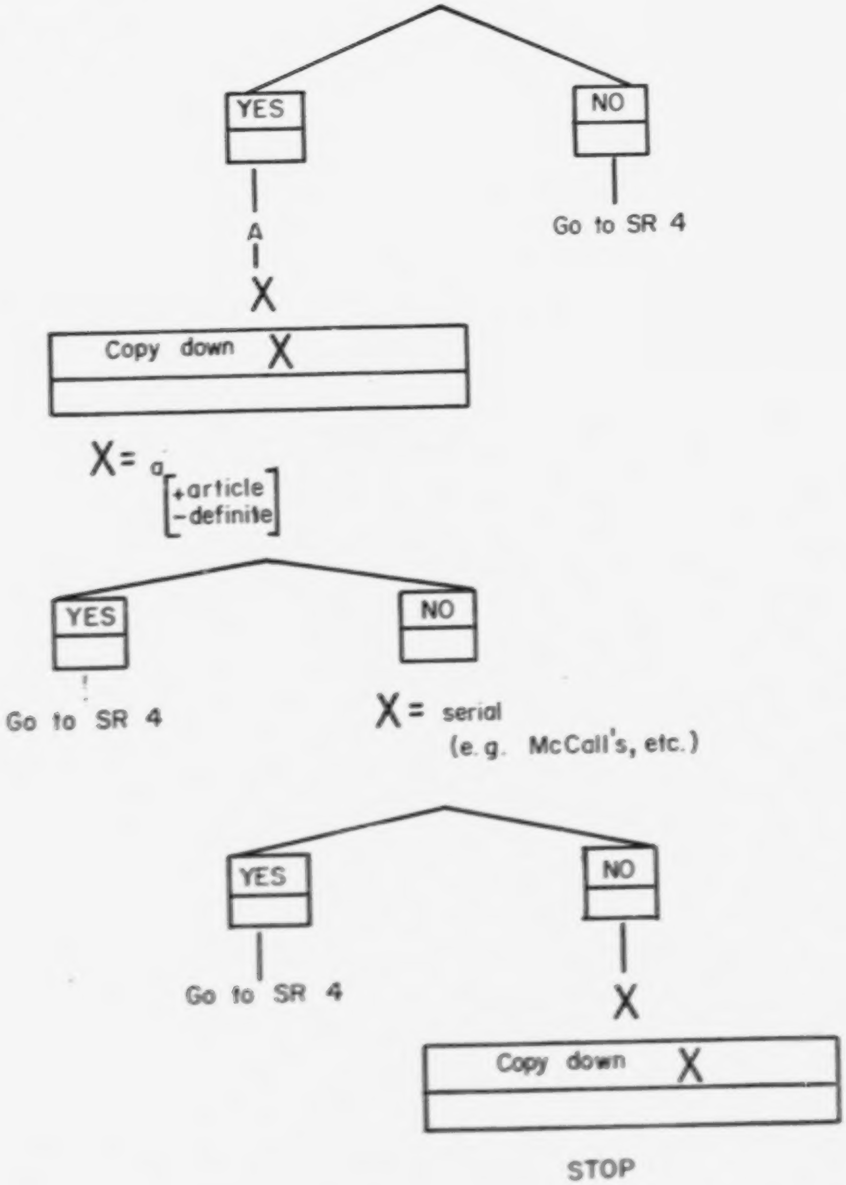
Where Z = anything or nothing but not BY


YES
NO
A
Go to SR 3
X
Copy down X
X= a
+ article
− definite
YES
NO
Go to SR 3
X= serial
(e.g. McCall's, etc.)
YES
NO
Go to SR 3
X
Copy down X
STOP

AASR 3
Recognition:

$$X_{\text{'s, s'}}\ Z \left[\begin{array}{l}\text{drawings}\\ \text{foreword}\\ \text{illustrat(ed) (ions)}\\ \text{introduc(ed) (tion)}\\ \text{to.....edited}\end{array}\right] Q\ {}_{\text{BY}}\ YQ$$

Where $Z =$ anything or nothing but not BY
$Q =$ anything or nothing

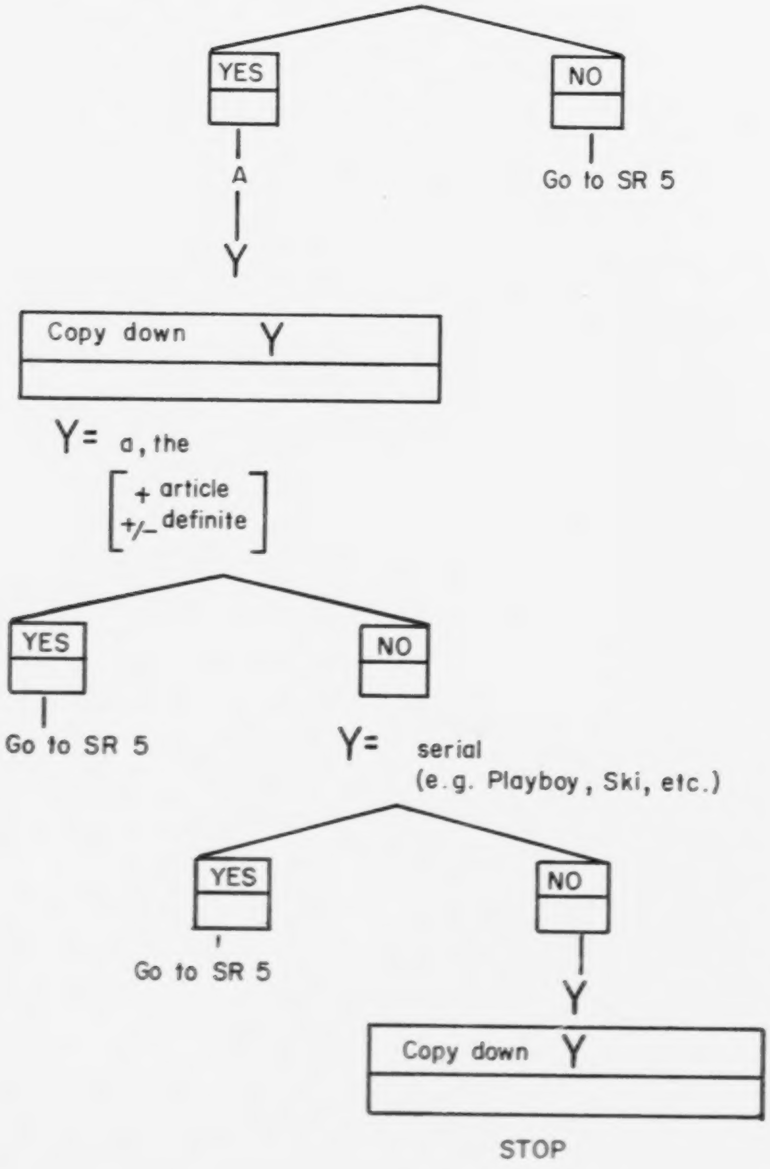

YES
NO
A
X
Go to SR 4
Copy down X
X = a [+article −definite]
YES
NO
Go to SR 4
X = serial
(e.g. McCall's, etc.)
YES
NO
Go to SR 4
X
Copy down X
STOP

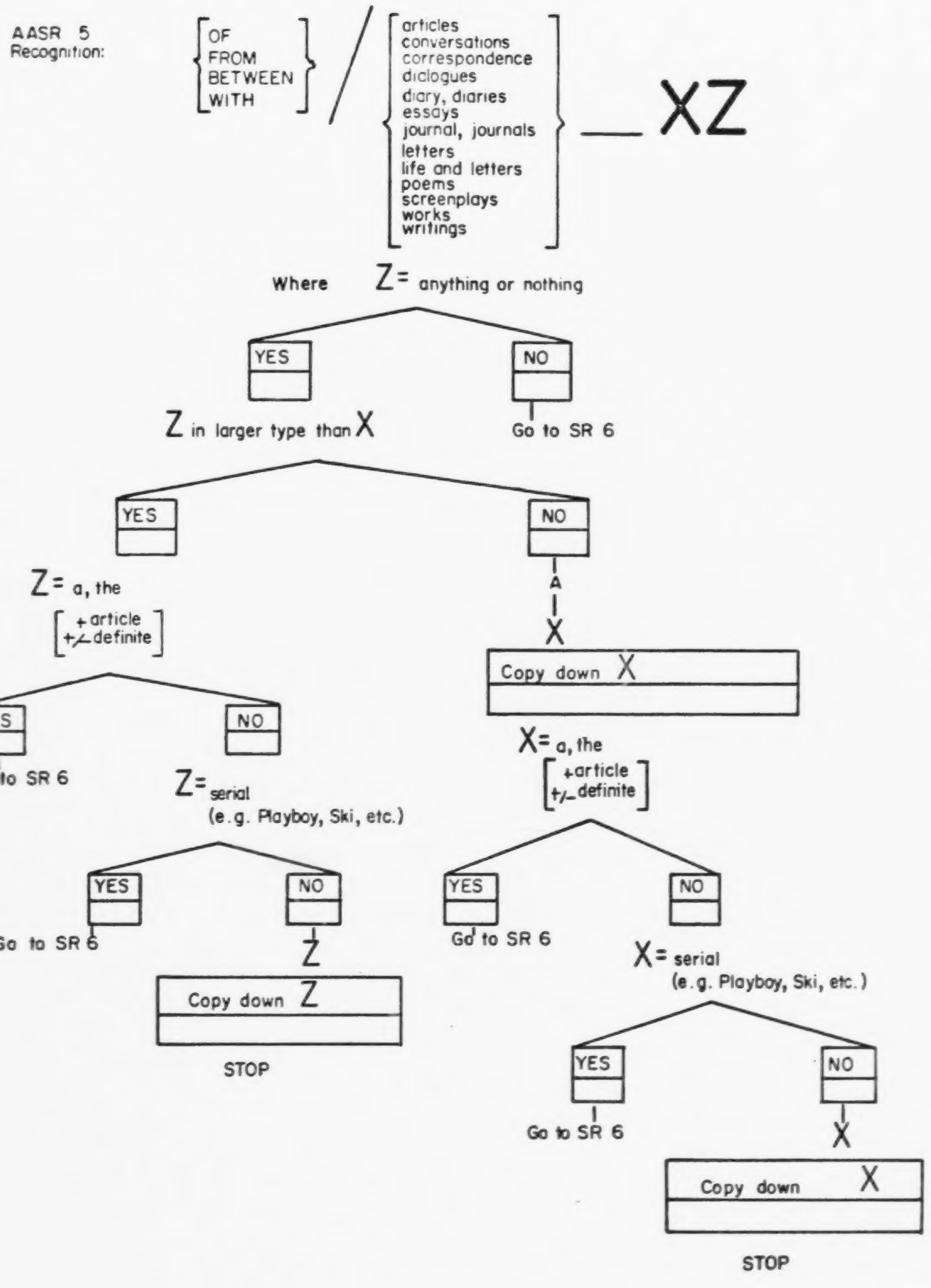
AASR 4
Recognition: OF/ Life and

- articles
- conversations
- correspondence
- dialogues
- diary, diaries
- essays
- journal, journals
- letters
- poems
- screenplays
- works
- writings

_XZ BY YZ

Where Z = anything or nothing but not EDITED before or after BY Y


YES
NO
A
Y
Go to SR 5
Copy down Y
Y= a, the
+ article
+/_ definite
YES
NO
Go to SR 5
Y= serial
(e.g. Playboy, Ski, etc.)
YES
NO
Go to SR 5
Y
Copy down Y
STOP

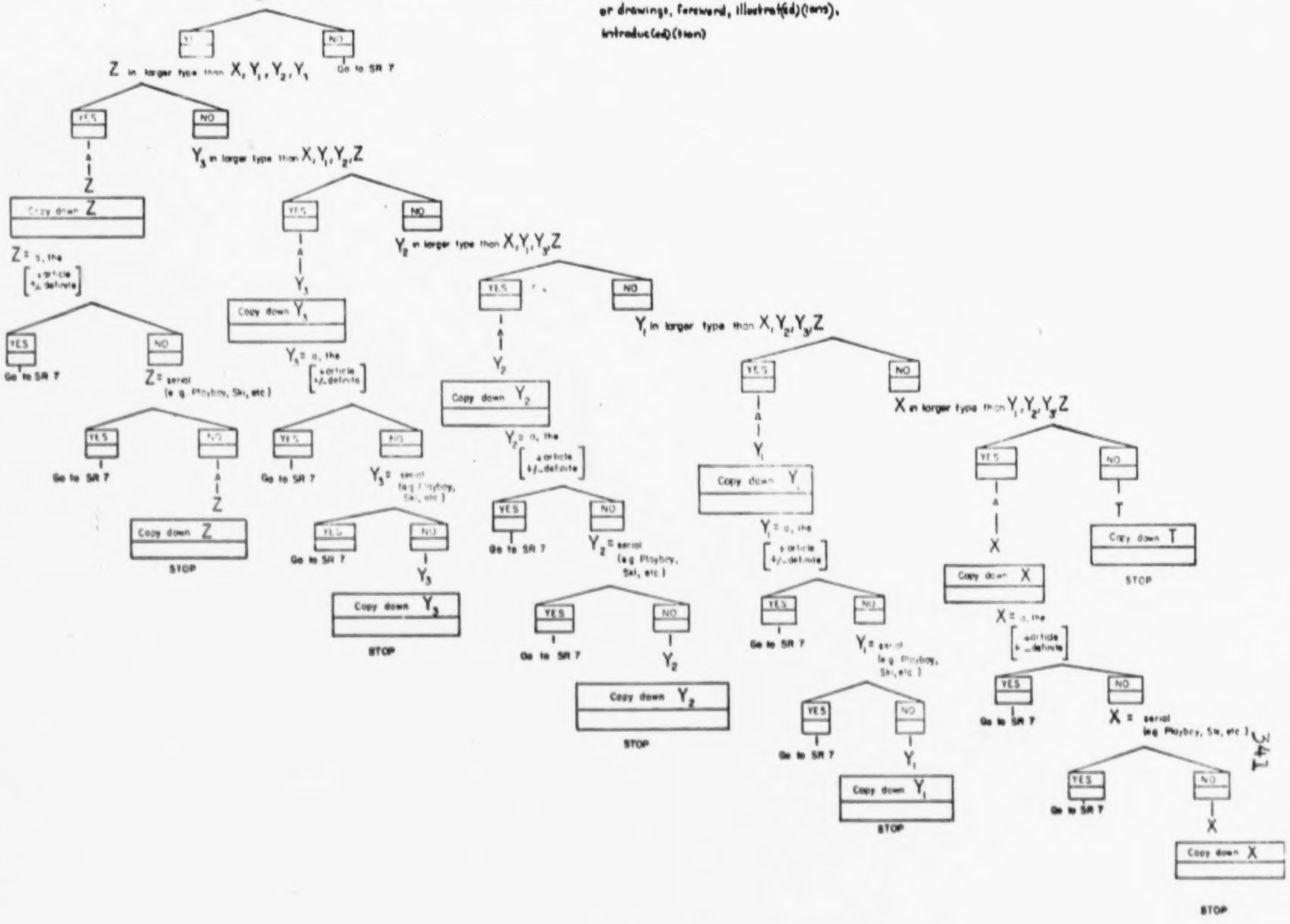

AASR 5
Recognition:
OF
FROM
BETWEEN
WITH
articles
conversations
correspondence
dialogues
diary, diaries
essays
journal, journals
letters
life and letters
poems
screenplays
works
writings
— XZ
Where Z= anything or nothing
YES
NO
Z in larger type than X
Go to SR 6
YES
NO
Z= a, the
+article
+/- definite
A
X
Copy down X
YES
NO
Go to SR 6
Z= serial
(e.g. Playboy, Ski, etc.)
X= a, the
+article
+/- definite
YES
NO
Go to SR 6
Z
Copy down Z
STOP
YES
NO
Go to SR 6
X= serial
(e.g. Playboy, Ski, etc.)
YES
NO
Go to SR 6
X
Copy down X
STOP

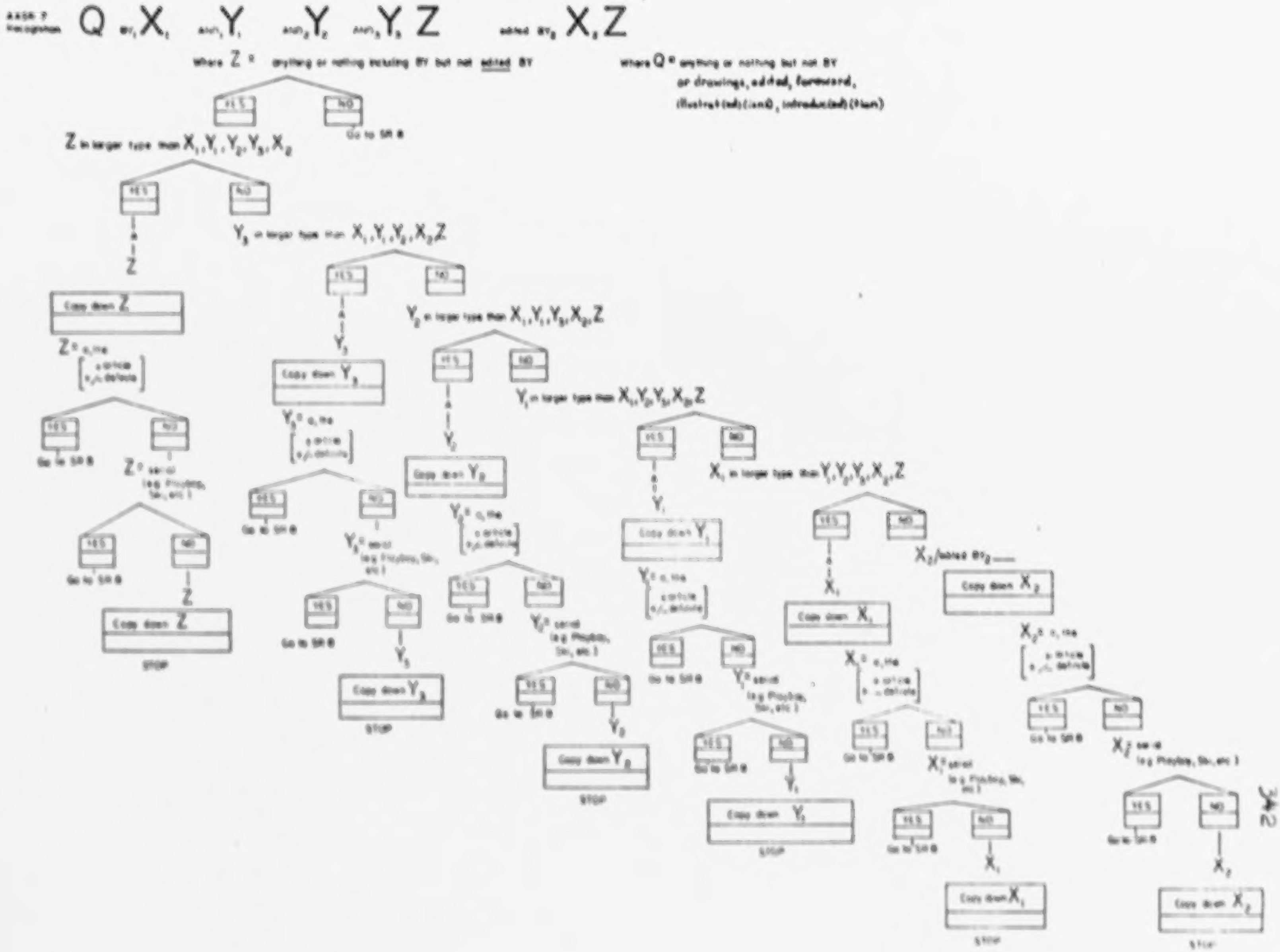
AACR 6
Recognition

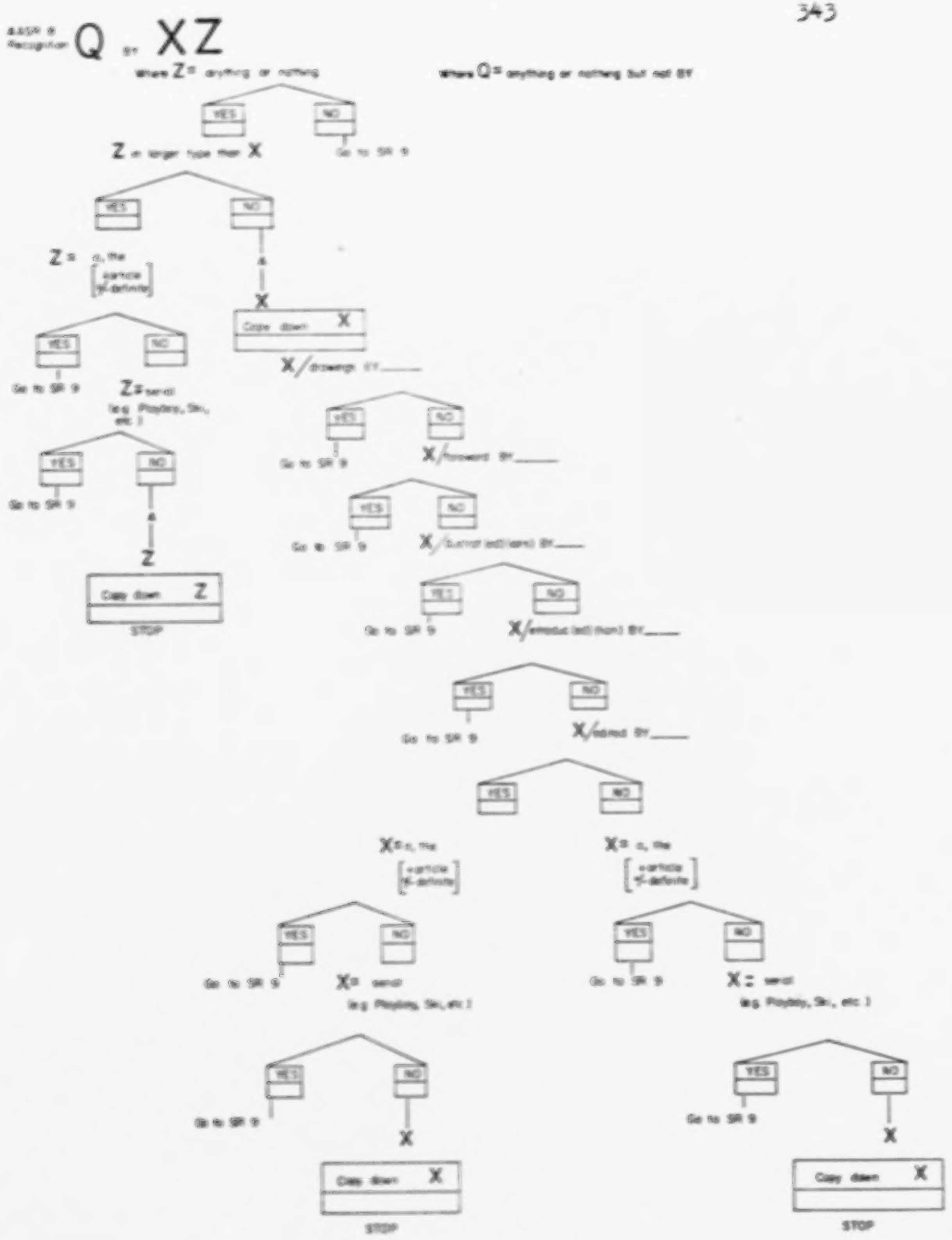
Q BY X AND$_1$ Y_1 AND$_2$ Y_2 AND$_3$ Y_3 Z

Where Z = anything or nothing but not BY

Where Q = anything or nothing but not BY
or drawings, foreword, illustrat(ed)(ions),
introduc(ed)(tion)


YES
NO
Go to SR 7
Z in larger type than X, Y1, Y2, Y1
YES
NO
Z
Copy down Z
Z = a, the
[+article +/-definite]
YES
Go to SR 7
NO
Z = serial (e.g. Playboy, Ski, etc)
YES
Go to SR 7
NO
Z
Copy down Z
STOP
Y3 in larger type than X, Y1, Y2, Z
YES
NO
Y3
Copy down Y3
Y3 = a, the
[+article +/-definite]
YES
Go to SR 7
NO
Y3 = serial (e.g. Playboy, Ski, etc.)
YES
Go to SR 7
NO
Y3
Copy down Y3
STOP
Y2 in larger type than X, Y1, Y3, Z
YES
NO
Y2
Copy down Y2
Y2 = a, the
[+article +/-definite]
YES
Go to SR 7
NO
Y2 = serial (e.g. Playboy, Ski, etc.)
YES
Go to SR 7
NO
Y2
Copy down Y2
STOP
Y1 in larger type than X, Y2, Y3, Z
YES
NO
Y1
Copy down Y1
Y1 = a, the
[+article +/-definite]
YES
Go to SR 7
NO
Y1 = serial (e.g. Playboy, Ski, etc.)
YES
Go to SR 7
NO
Y1
Copy down Y1
STOP
X in larger type than Y1, Y2, Y3, Z
YES
NO
X
Copy down X
X = a, the
[+article +/-definite]
YES
Go to SR 7
NO
X = serial (e.g. Playboy, Ski, etc.)
YES
Go to SR 7
NO
X
Copy down X
STOP
T
Copy down T
STOP


Q BY₁ X₁ AND₁ Y₁ AND₂ Y₂ AND₃ Y₃ Z edited BY₂ X₂ Z
Where Z = anything or nothing including BY but not edited BY
Where Q = anything or nothing but not BY or drawings, edited, foreword, illustrat(ed)(ions), introduc(ed)(tion)
YES
NO
Go to SR 8
Z in larger type than X₁, Y₁, Y₂, Y₃, X₂
Y₃ in larger type than X₁, Y₁, Y₂, X₂, Z
Y₂ in larger type than X₁, Y₁, Y₃, X₂, Z
Y₁ in larger type than X₁, Y₂, Y₃, X₂, Z
X₁ in larger type than Y₁, Y₂, Y₃, X₂, Z
X₂ / edited BY₂
Copy down Z
Copy down Y₃
Copy down Y₂
Copy down Y₁
Copy down X₁
Copy down X₂
Z = a, the [a article / a, b definite]
Z = serial (e.g. Playboy, Ski, etc.)
STOP


AASR 8
Recognition
Q BY XZ
Where Z = anything or nothing
Where Q = anything or nothing but not BY
YES
NO
Z in larger type than X
Go to SR 9
Z = a, the [+article +/- definite]
&
X
Copy down X
X/drawings BY____
Z ≠ serial (e.g. Playboy, Ski, etc.)
Go to SR 9
X/foreword BY____
Z
Copy down Z
STOP
X/[illegible] (editions) BY____
X/introduction (edition) BY____
X/edited BY____
X ≠ a, the [+article +/- definite]
X = a, the [+article +/- definite]
X ≠ serial (e.g. Playboy, Ski, etc.)
X = serial (e.g. Playboy, Ski, etc.)
Copy down X
STOP

AASR 9
Recognition:

Everything else ——→ T

Where Everything else = recognitions to which previous rules did not apply

——→ = goes to

T = Title

T

Copy down T

STOP

AA Selection Rules

SR 1 $X \overset{\&}{\text{AND}} Y \text{ 'S,S' } Z \text{ BY } W Z \longrightarrow \text{A}$

Where X, Y, W = author names
Z = anything or nothing

And where A = $X \neq$ a $\begin{bmatrix} + \text{article} \\ - \text{definite} \end{bmatrix}$

and $\neq$ a serial
(e.g. Field & Stream's, etc.)

SR 2 $X \text{ 'S,S' } Z \longrightarrow \text{A}$

Where X = an author name
Z = anything or nothing but not BY

And where A = $X \neq$ a $\begin{bmatrix} + \text{article} \\ - \text{definite} \end{bmatrix}$

and $\neq$ a serial
(e.g. McCall's, etc.)

SR 3 $X \text{ 'S,S' } Z \left\{ \begin{matrix} \text{drawings} \\ \text{foreword} \\ \text{illustrat(ed)(ions)} \\ \text{introduc(ed)(tion)} \\ \text{to...edited} \end{matrix} \right\} Q \text{ BY } Y Q$

Where X, Y = author names
Z = anything or nothing but not BY
Q = anything or nothing

And where A = $X \neq$ a $\begin{bmatrix} + \text{article} \\ - \text{definite} \end{bmatrix}$

and $\neq$ a serial
(e.g. McCall's, etc.)

$$\text{SR 4 OF} \,/\, \text{Life and} \left\{\begin{array}{l} \text{articles}\\ \text{conversations}\\ \text{correspondence}\\ \text{dialogues}\\ \text{diary, diaries}\\ \text{essays}\\ \text{journal, journals}\\ \text{letters}\\ \text{poems}\\ \text{works}\\ \text{writings} \end{array}\right\} \_\, XZ \text{ BY } YZ \longrightarrow A$$

Where X, Y = author names
Z = nothing or anything but not EDITED before or after BY Y

And where $A = Y \neq$ a, the $\left[\begin{array}{l} +\ \text{article}\\ +/-\ \text{definite} \end{array}\right]$

and $\neq$ a serial (e.g. Playboy, Ski, etc.)

$$\text{SR 5} \left\{\begin{array}{l} \text{OF}\\ \text{FROM}\\ \text{BETWEEN}\\ \text{WITH} \end{array}\right\} \Big/ \left\{\begin{array}{l} \text{articles}\\ \text{conversations}\\ \text{correspondence}\\ \text{dialogues}\\ \text{diary, diaries}\\ \text{essays}\\ \text{journal, journals}\\ \text{letters}\\ \text{poems}\\ \text{works}\\ \text{writings} \end{array}\right\} \_\, XZ \longrightarrow A$$

Where X = an author name
Z = anything or nothing

And where $A = X \neq$ a, the $\left[\begin{array}{l} +\ \text{article}\\ +/-\ \text{definite} \end{array}\right]$

and $\neq$ a serial (e.g. Playboy, Ski, etc.)

SR 6 $Q_{\text{BY}}\, X_{\text{AND}_1}\, Y_1 \quad \text{AND}_2\, Y_2 \quad \text{AND}_3\, Y_3\, Z \longrightarrow T$

Where X, Y = author names
Z = anything or nothing but not BY
Q = anything or nothing but not BY or drawings, foreword, illustrat(ed)(ions), introduc(ed)(tion)

And where T = title

Unless Z or X or Y_1, Y_2, Y_3 is in larger type

and ≠ a, the $\begin{bmatrix} + \text{article} \\ +/- \text{definite} \end{bmatrix}$

and ≠ a serial
(e.g. Playboy, Ski, etc.)

SR 7 $\text{BY}_1\, X_1 \quad \text{AND}_1\, Y_1 \quad \text{AND}_2\, Y_2 \quad \text{AND}_3\, Y_3 \quad \text{edited BY}_2\, X_2\, Z \rightarrow A$

Where X, Y = author names
Z = anything or nothing including BY but not edited BY
Q = anything or nothing but not BY or drawings, edited, foreword, illustrat(ed)(ions, introduc(ed)(tion)

And where A = X /edited BY___

Where X ≠ a, the $\begin{bmatrix} + \text{article} \\ +/- \text{definite} \end{bmatrix}$

and ≠ a serial
(e.g. Playboy, Ski, etc.)

Unless Z or X_1 or Y_1, Y_2, Y_3 is in larger type

and ≠ a, the $\begin{bmatrix} + \text{article} \\ +/- \text{definite} \end{bmatrix}$

and ≠ a serial
(e.g. Playboy, Ski, etc.)

SR 8 $Q\,{\scriptstyle BY}\,X Z \longrightarrow A$

With X = author name
Z = anything or nothing
Q = anything or nothing but not BY

And where $A = X$

Where $X \neq / \left\{ \begin{array}{l} \text{drawings} \\ \text{foreword} \\ \text{illustrat(ed)(ions)} \\ \text{introduc(ed)(tion)} \end{array} \right\}$ BY___

Where $X \neq$ a, the $\begin{bmatrix} +\text{ article} \\ +/-\text{ definite} \end{bmatrix}$

and $\neq$ a serial
(e.g. Playboy, Ski, etc.)

Unless Z is in larger type

and $\neq$ a, the $\begin{bmatrix} +\text{ article} \\ +/-\text{ definite} \end{bmatrix}$

and $\neq$ a serial
(e.g. Playboy, Ski, etc.)

SR 9 Everything else $\longrightarrow$ T

Where T = title

ALA Recognition Rules

RR 1(a) Recognize BY in any substring

Recognition YES ☐ NO ☐

If YES, copy down recognition for the rule, set off with appropriate boundary markers in the appropriate string

____

____

Go to RR 1(b)

If NO, go to RR 1(b)

RR 1(b) Recognize BY as implied at the head of any substring in which there is no physical expression of the marker in it

Recognition YES ☐ NO ☐

If YES, copy down recognition for the rule with BY in brackets, [ ], set off with appropriate boundary markers in the appropriate string

____

____

Go to RR 1(c)

If NO, go to RR 1(c)

RR 1(c) Recognize $\left\{\begin{array}{l}\text{drawings}\\ \text{edited}\\ \text{foreword}\\ \text{illustrat(ed)(ions)}\\ \text{introduc(ed)(tion)}\end{array}\right\}$ / ____ BY

In any substring

Recognition YES ☐ NO ☐

If $\boxed{\text{YES}}$, copy down recognition for the rule and carry recognition for the previous rule forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 2(a)

If $\boxed{\text{NO}}$, go to RR 2(a)

RR 2(a) Recognize (A) ___$\overset{\&}{\text{AND}}$___'S,S' when

(A)___$\overset{\&}{\text{AND}}$___'S,S' / ___...BY [BY] ...

Recognition $\boxed{\text{YES}}$ $\boxed{\text{NO}}$

If $\boxed{\text{YES}}$, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 2(b)

If $\boxed{\text{NO}}$, go to RR 2(b)

RR 2(b) Recognize (A)___'S,S' when

$$\text{(A)\_\_\_'S,S'} \;/\; \text{\_\_\_...} \left\{\begin{array}{l}\text{drawings}\\ \text{foreword}\\ \text{illustrat(ed)(ions)}\\ \text{introduc(ed)(tion)}\\ \text{to...edited}\end{array}\right\} \begin{array}{l}\text{...BY}\\ \text{[BY]...}\end{array}$$

Recognition $\boxed{\text{YES}}$ $\boxed{\text{NO}}$

If YES, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 2(c)

If NO, go to RR 2(c)

RR 2(c) Recognize (A)___'S,S' when it occurs without BY [BY] in the substring

Recognition YES ☐ NO ☐

If YES, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 3

If NO, go to RR 3

RR 3 Recognize OF when

OF / Life and { articles
conversations
correspondence
dialogues
diary, diaries
essays
journal, journals
letters
poems
screenplays
works
writings } ___...

Recognition YES ☐ NO ☐

If YES , copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 4

If NO , go to RR 4

RR 4 Recognize OF, FROM, BETWEEN, WITH when

$$\left\{\begin{array}{l}\text{OF}\\ \text{FROM}\\ \text{BETWEEN}\\ \text{WITH}\end{array}\right\} \Big/ \left\{\begin{array}{l}\text{articles}\\ \text{conversations}\\ \text{correspondence}\\ \text{dialogues}\\ \text{diary, diaries}\\ \text{essays}\\ \text{journal, journals}\\ \text{letters}\\ \text{poems}\\ \text{screen plays}\\ \text{works}\\ \text{writings}\end{array}\right\} \_\_\ldots$$

Recognition YES NO

If YES , copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 5

If NO , go to RR 5

RR 5 Recognize OF after BY in the substring

Recognition

YES	NO

If YES, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 6(a)

If NO, go to RR 6(a)

RR 6(a) Recognize AND as implied at the end of each string except the final string in the substring when:

(a) BY [BY] is at the head of the substring and there is no other physical expression of the marker BY or the markers AND, OF, or WITH in the substring

(b) BY [BY] is at the head of the substring and EDITED BY is at the end of the substring and there is no other physical expression of the marker BY or the markers AND, OF, or WITH in the substring

Recognition

YES	NO

If YES, copy down recognition for the rule with AND in brackets, [ ], and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 6(b)

If NO, go to RR 6(b)

RR 6(b) Recognize AND, as implied in the form of a comma (,) or slash (/) in the substring, when:

(a) BY [BY] is at the head of the substring and there is no other physical expression of the marker BY or the markers OF or WITH in the substring

(b) BY [BY] is at the head of the substring and EDITED BY is at the end of the substring and there is no other physical expression of the marker BY or the markers OF or WITH in the substring

Recognition YES ☐ NO ☐

If YES, copy down recognition for the rule with AND in brackets, , and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 6(c)

If NO, go to RR 6(c)

RR 6(c) Recognize AND, &, after BY in the substring, except when:

(a) $\left\{\begin{matrix}\text{AND}\\ \&\end{matrix}\right\} \Big/ \text{BY}[\text{BY}]\ldots\text{WITH}\ldots\_\_\_$

(b) $\left\{\begin{matrix}\text{AND}\\ \&\\ \text{AND}\end{matrix}\right\} \Big/ \text{BY}[\text{BY}]\ldots\left\{\begin{matrix}\text{edited}\_\_\_\\ \_\_\_\text{edited}\end{matrix}\right\}\ldots$

Recognition YES ☐ NO ☐

If YES, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 7

If $\boxed{\text{NO}}$, go to RR 7

RR 7 Recognize WITH after BY in the substring except when

(a) WITH / BY [BY]...___...BY [BY]

(b) WITH / BY [BY]...___...OF

Recognition	YES	NO

If $\boxed{\text{YES}}$, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

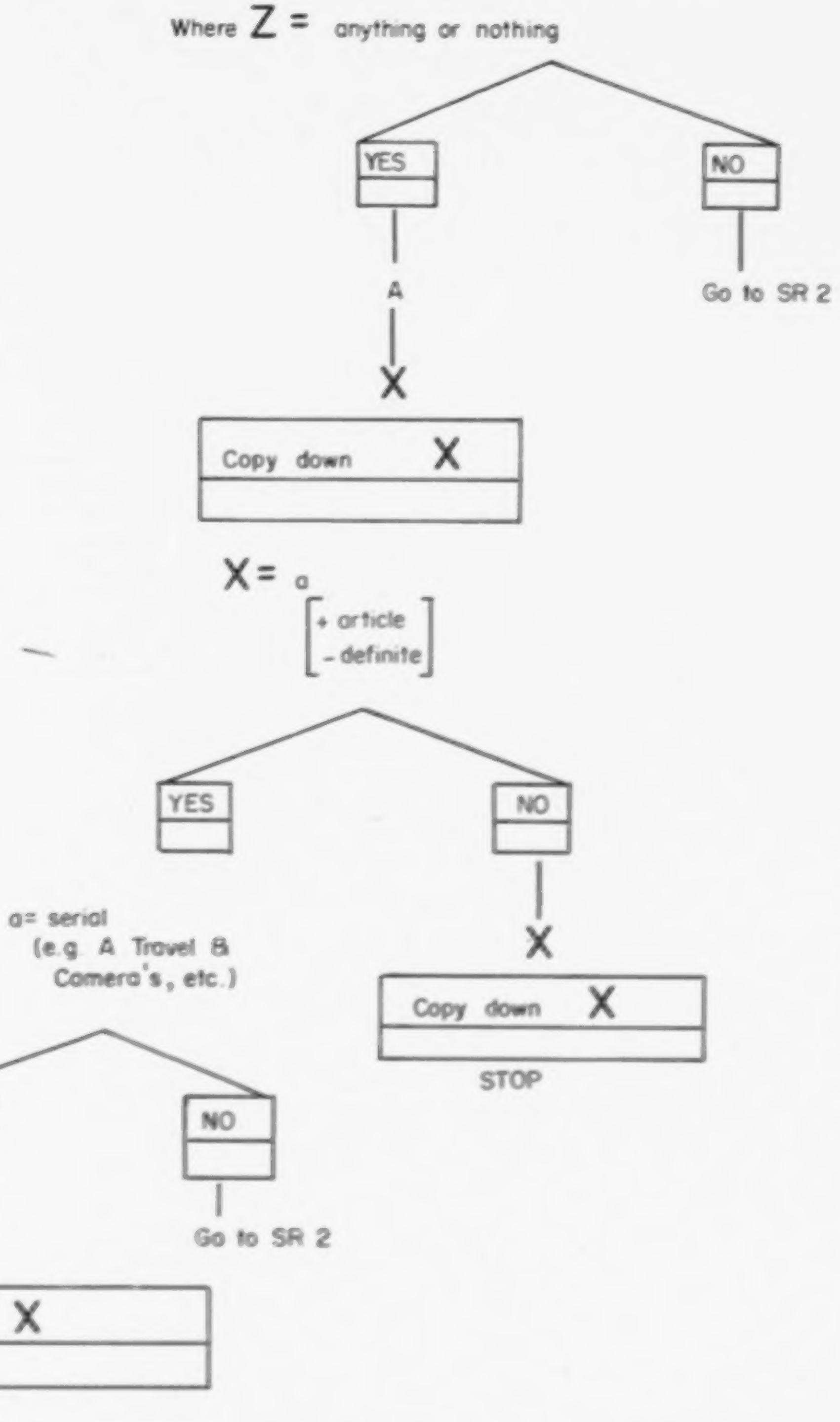
____

Apply ALA Selection Rules to the recognition

If $\boxed{\text{NO}}$, apply ALA Selection Rules to the last recognition copied down under the rules above or to the "non" or no recognition as the case may be

ALASR I
Recognition

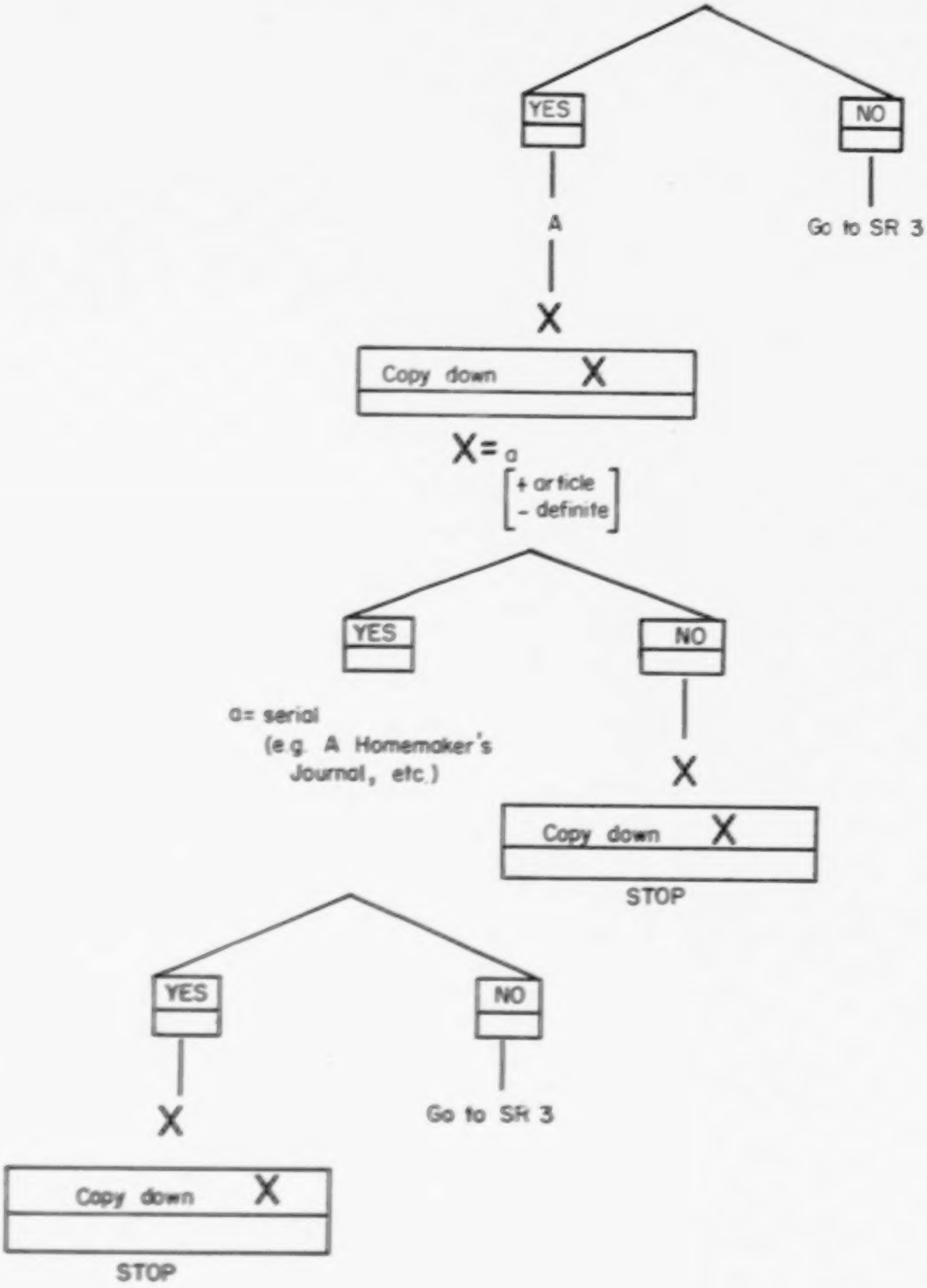
X & AND Y 's, s' Z BY WZ


Where Z = anything or nothing
YES
NO
A
Go to SR 2
X
Copy down X
X = a
+ article
- definite
YES
NO
a= serial
(e.g. A Travel &
Camera's, etc.)
X
Copy down X
STOP
YES
NO
X
Go to SR 2
Copy down X
STOP

ALASR 2
Recognition:

$X_{\text{'s, s'}}\ Z$

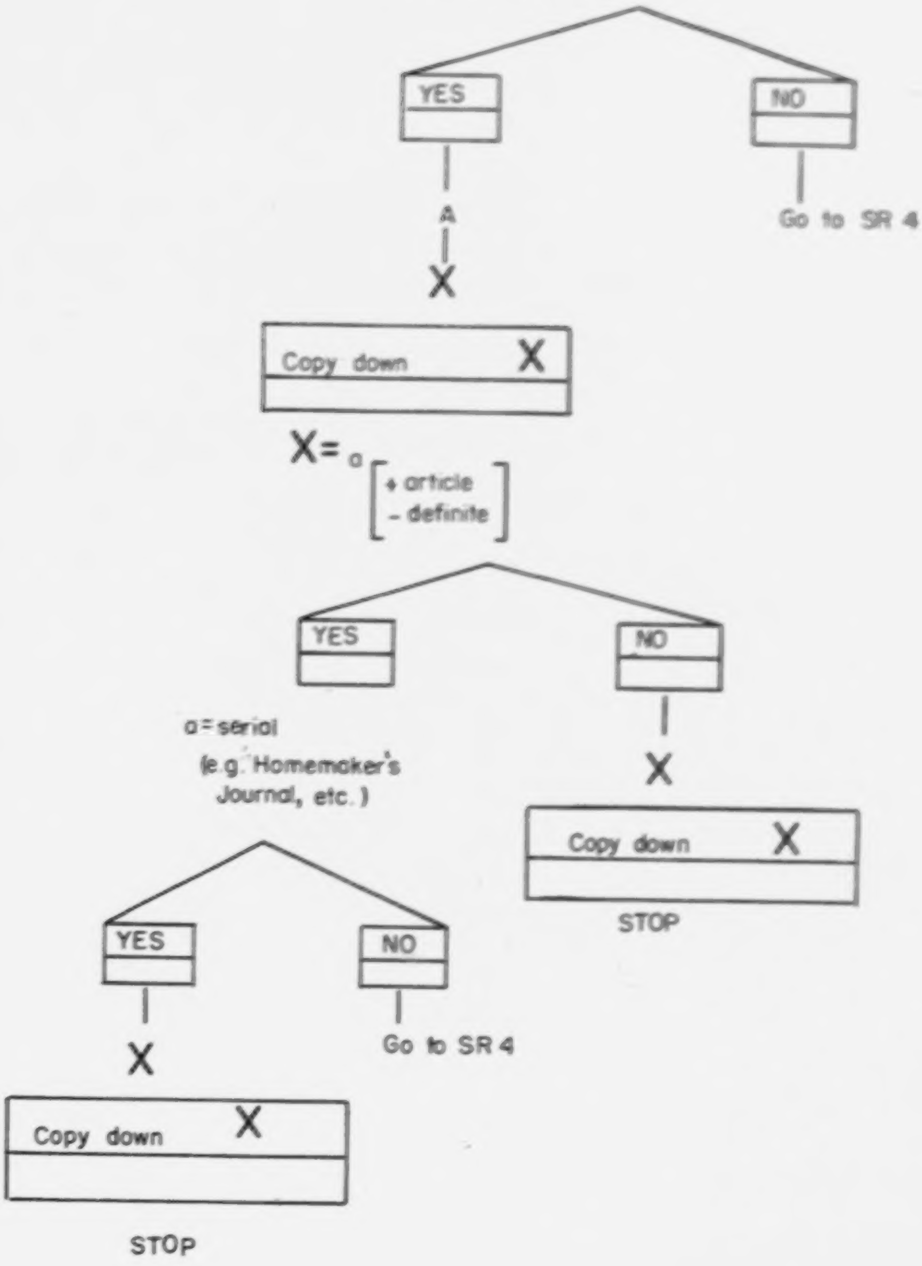
Where $Z=$ anything or nothing but not BY

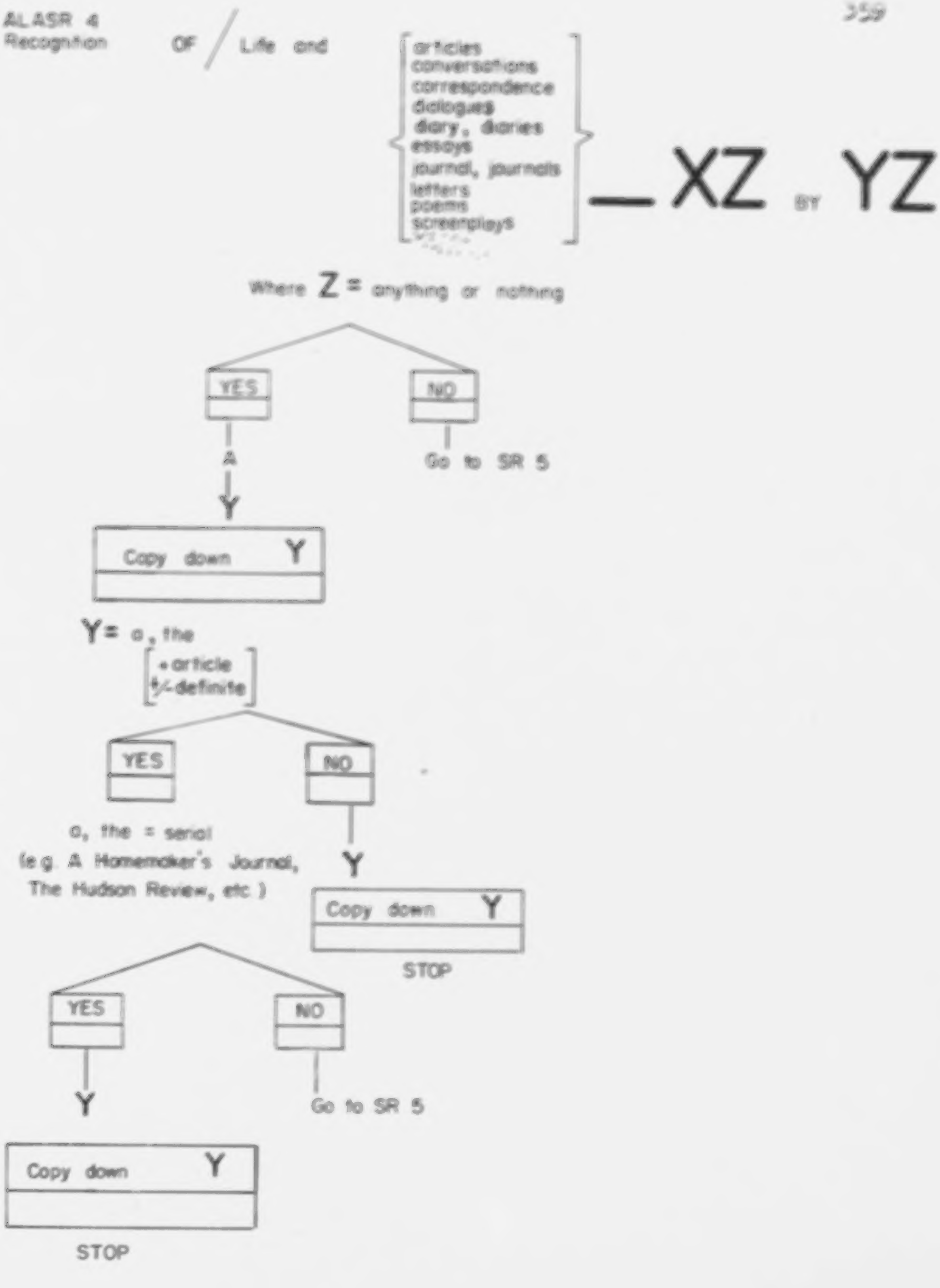

YES
NO
A
Go to SR 3
X
Copy down X
X= a
[+ article
- definite]
YES
NO
a= serial
(e.g. A Homemaker's
Journal, etc.)
X
Copy down X
STOP
YES
NO
X
Go to SR 3
Copy down X
STOP

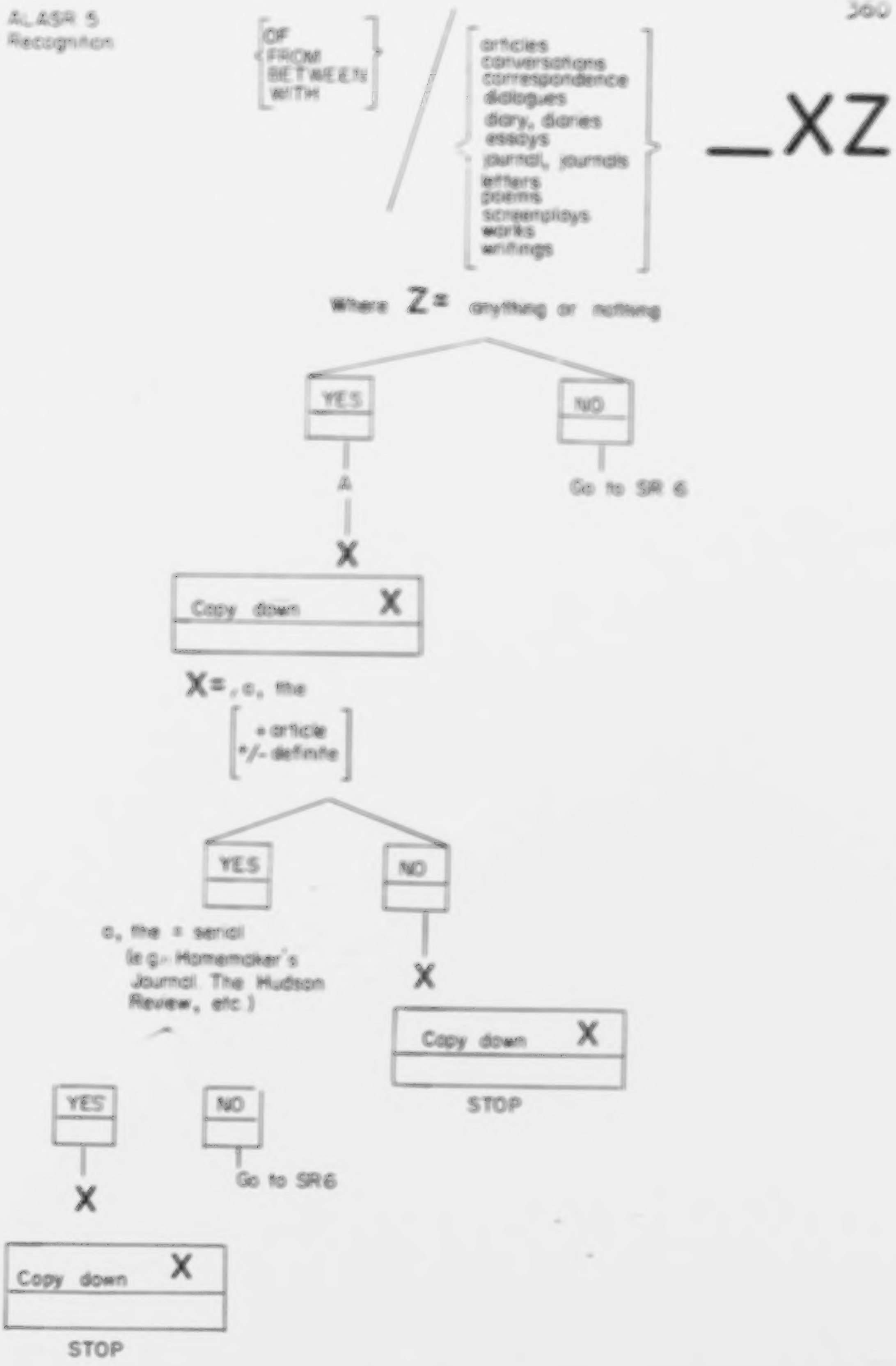
ALASR 3
Recognition: X 's, s' Z
{ drawings
foreword
illustrat (ed) (ions)
introduc (ed) (tion)
to . . . edited }
Q BY YQ

Where Z= anything or nothing but not BY

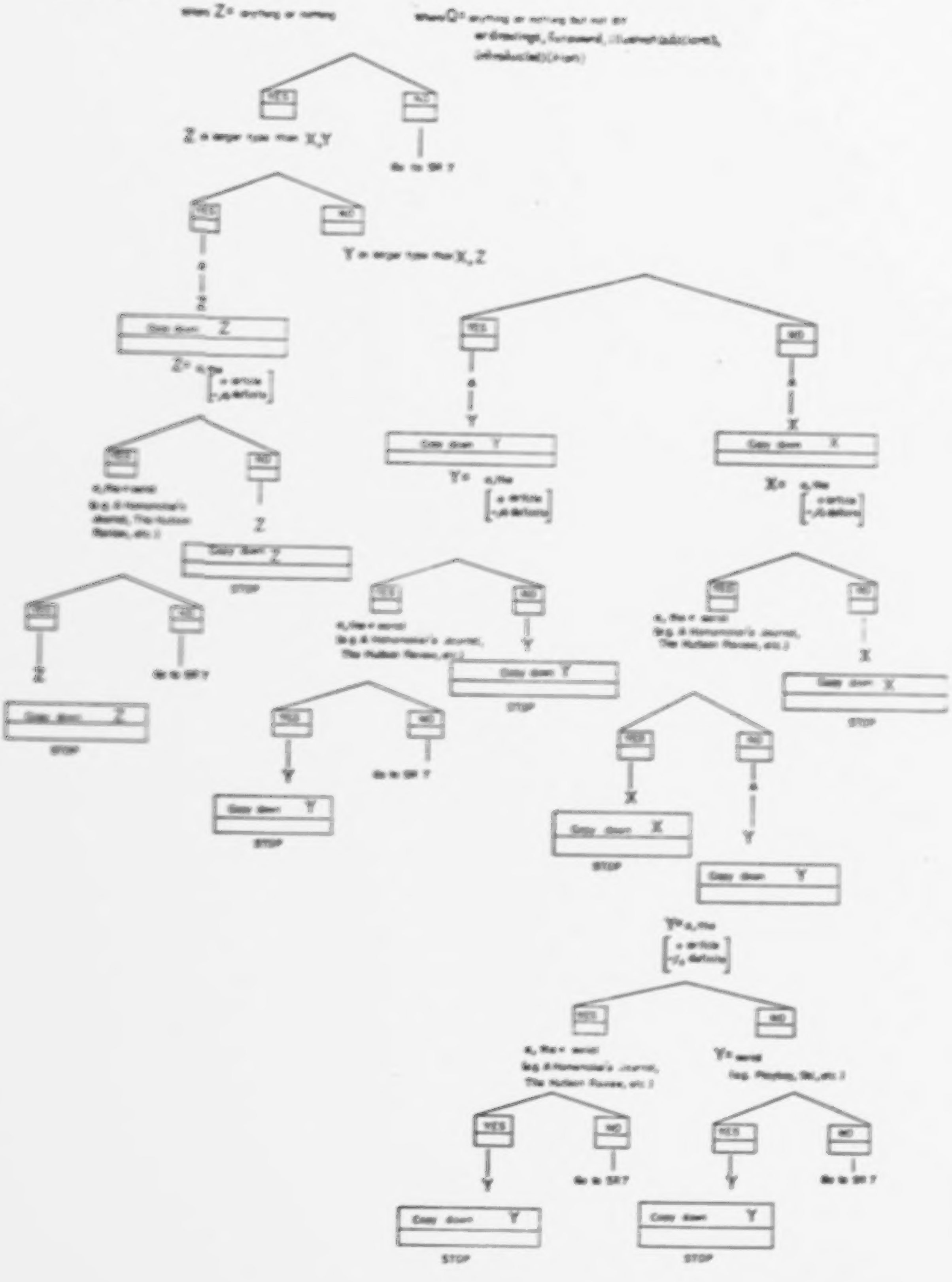
Q= anything or nothing

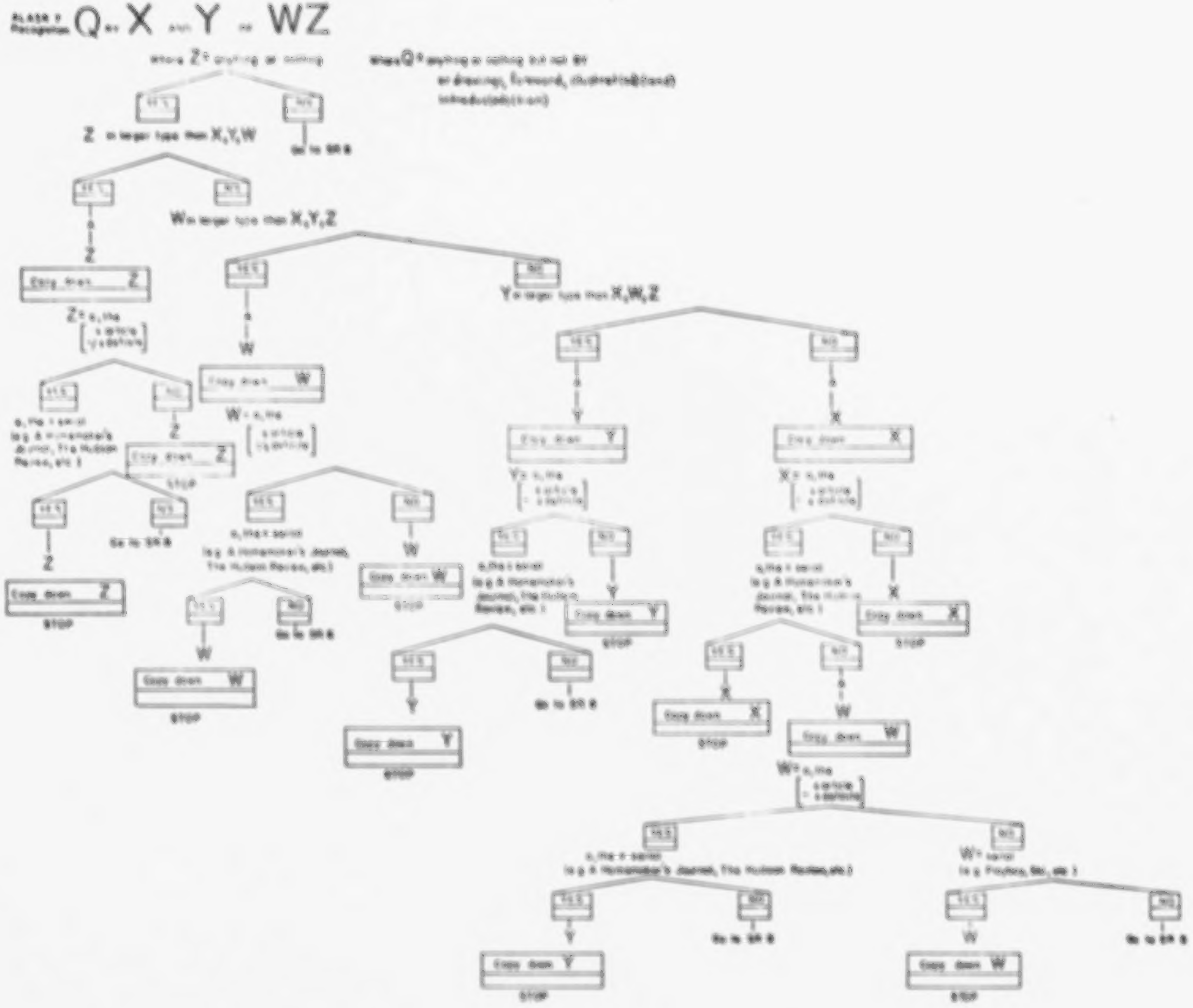

YES
NO
A
X
Go to SR 4
Copy down X
X= a [+ article - definite]
YES
NO
a=serial (e.g. Homemaker's Journal, etc.)
X
Copy down X
STOP
YES
NO
X
Go to SR 4
Copy down X
STOP

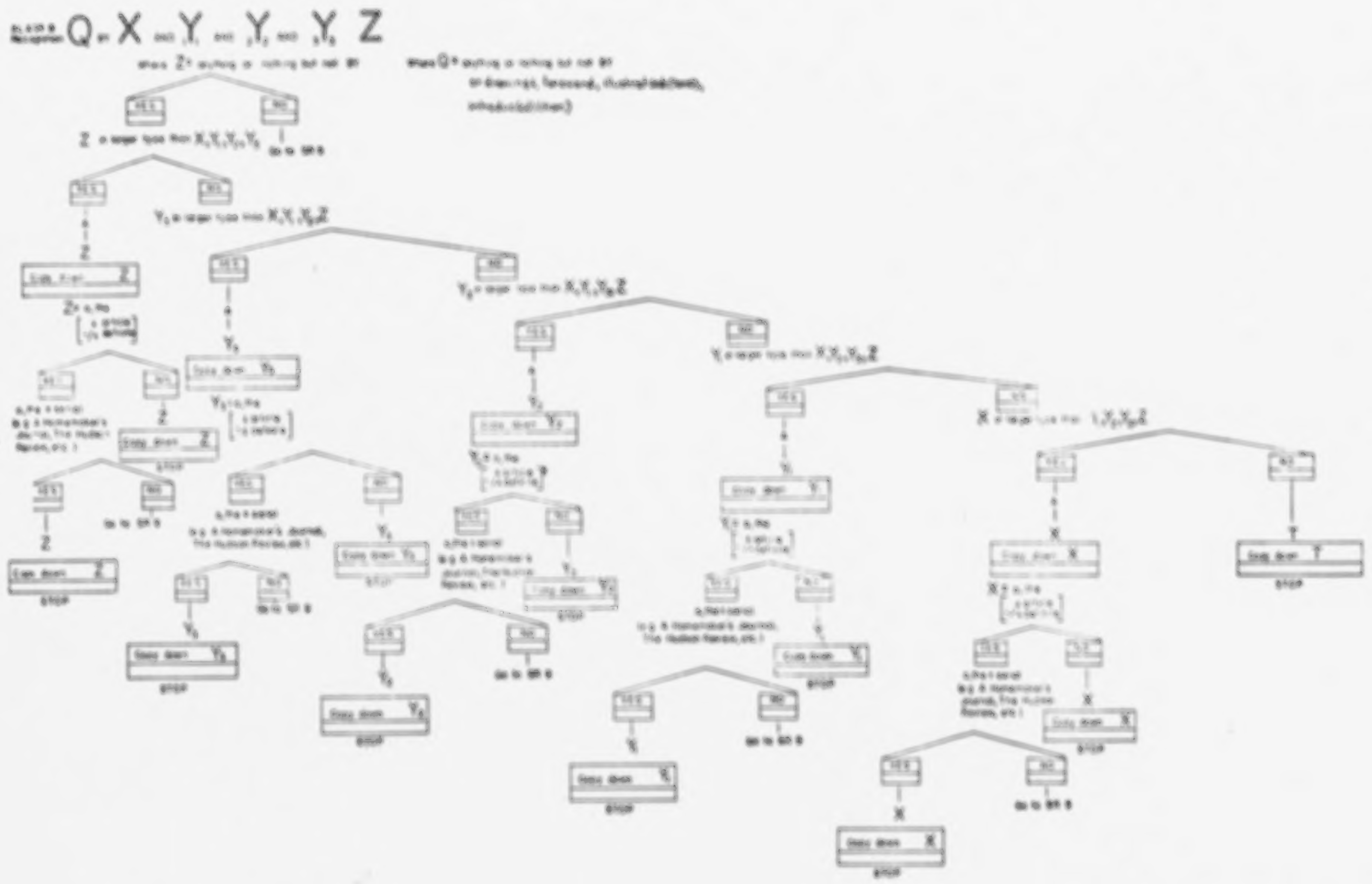

ALASR 4
Recognition
OF / Life and
articles
conversations
correspondence
dialogues
diary, diaries
essays
journal, journals
letters
poems
screenplays
— XZ BY YZ
Where Z = anything or nothing
YES
NO
A
Y
Go to SR 5
Copy down Y
Y= a, the
+ article
+/- definite
YES
NO
a, the = serial
(e.g. A Homemaker's Journal,
The Hudson Review, etc.)
Y
Copy down Y
STOP
YES
NO
Y
Go to SR 5
Copy down Y
STOP


OF
FROM
BETWEEN
WITH
articles
conversations
correspondence
dialogues
diary, diaries
essays
journal, journals
letters
poems
screenplays
works
writings
_XZ
Where Z = anything or nothing
YES
NO
A
Go to SR 6
X
Copy down X
X = , a, the
+ article
+/- definite
YES
NO
a, the = serial
(e.g., Homemaker's Journal, The Hudson Review, etc.)
X
Copy down X
STOP
YES
NO
X
Go to SR6
Copy down X
STOP

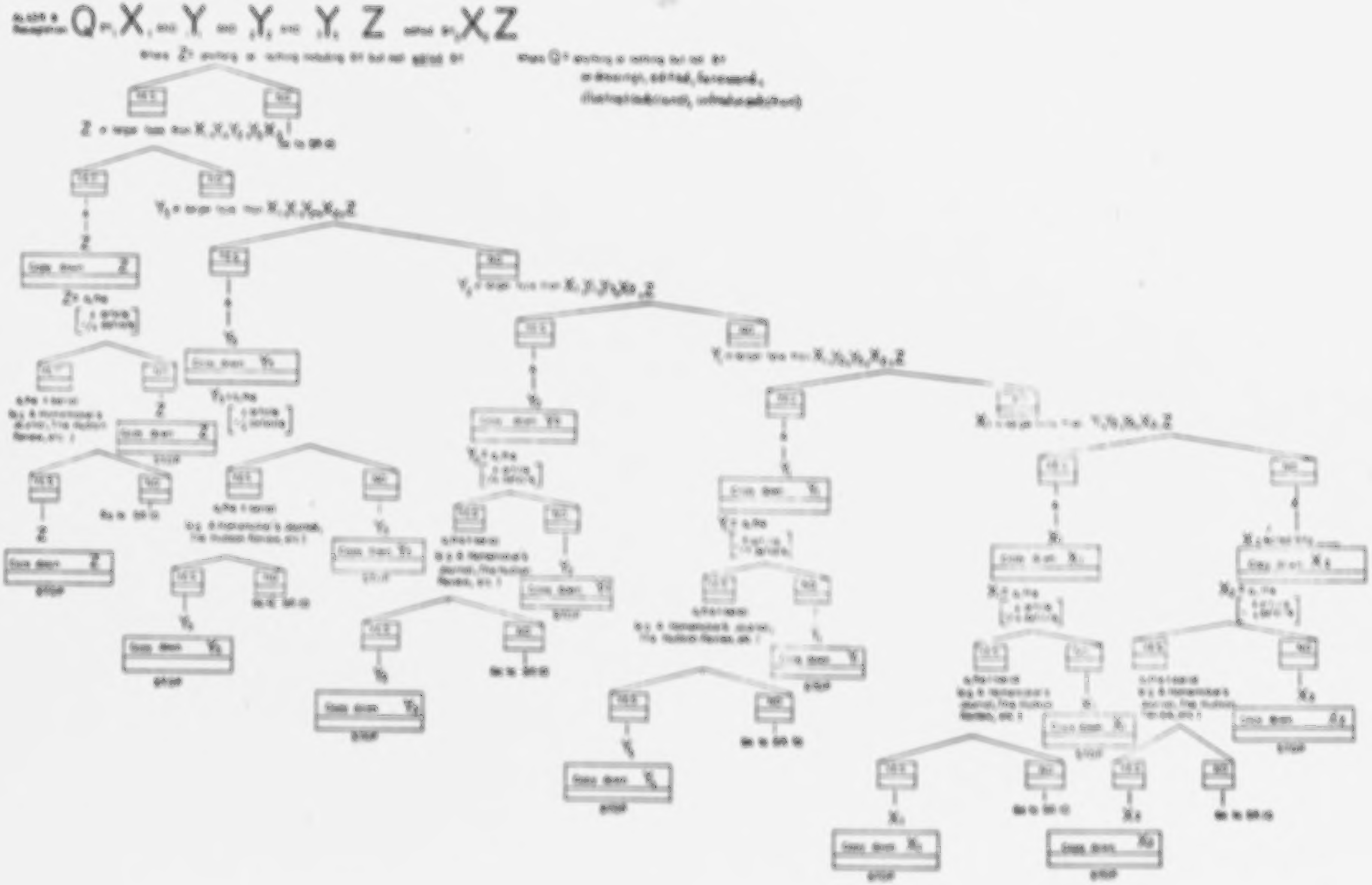
Q - X or YZ


where Z = anything or nothing
where Q = anything or nothing but not ill. or drawings, foreword, illustrat(ed)(ions), introduc(ed)(tion)
YES
NO
Z is larger type than X, Y
Go to SR 7
Y is larger type than X, Z
Copy down Z
Z = a, the [= article / -, ø definite]
a, the + serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)
STOP
Copy down Y
Y = a, the [= article / -, ø definite]
Copy down X
X = a, the [= article / -, ø definite]
Y = serial (e.g. Playboy, Ski, etc.)


Q ... X ... Y ... WZ
where Z = anything or nothing
where Q = anything or nothing but not (in) or drawings, foreword, illustrat(ed)(ions) introduc(ed)(tion)
YES
NO
Z in larger type than X, Y, W
Go to SR B
W in larger type than X, Y, Z
Y in larger type than X, W, Z
Copy down Z
Copy down W
Copy down Y
Copy down X
Z = a, the
W = a, the
Y = a, the
X = a, the
a, the + serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)
W = serial (e.g. Playboy, Ski, etc.)
STOP

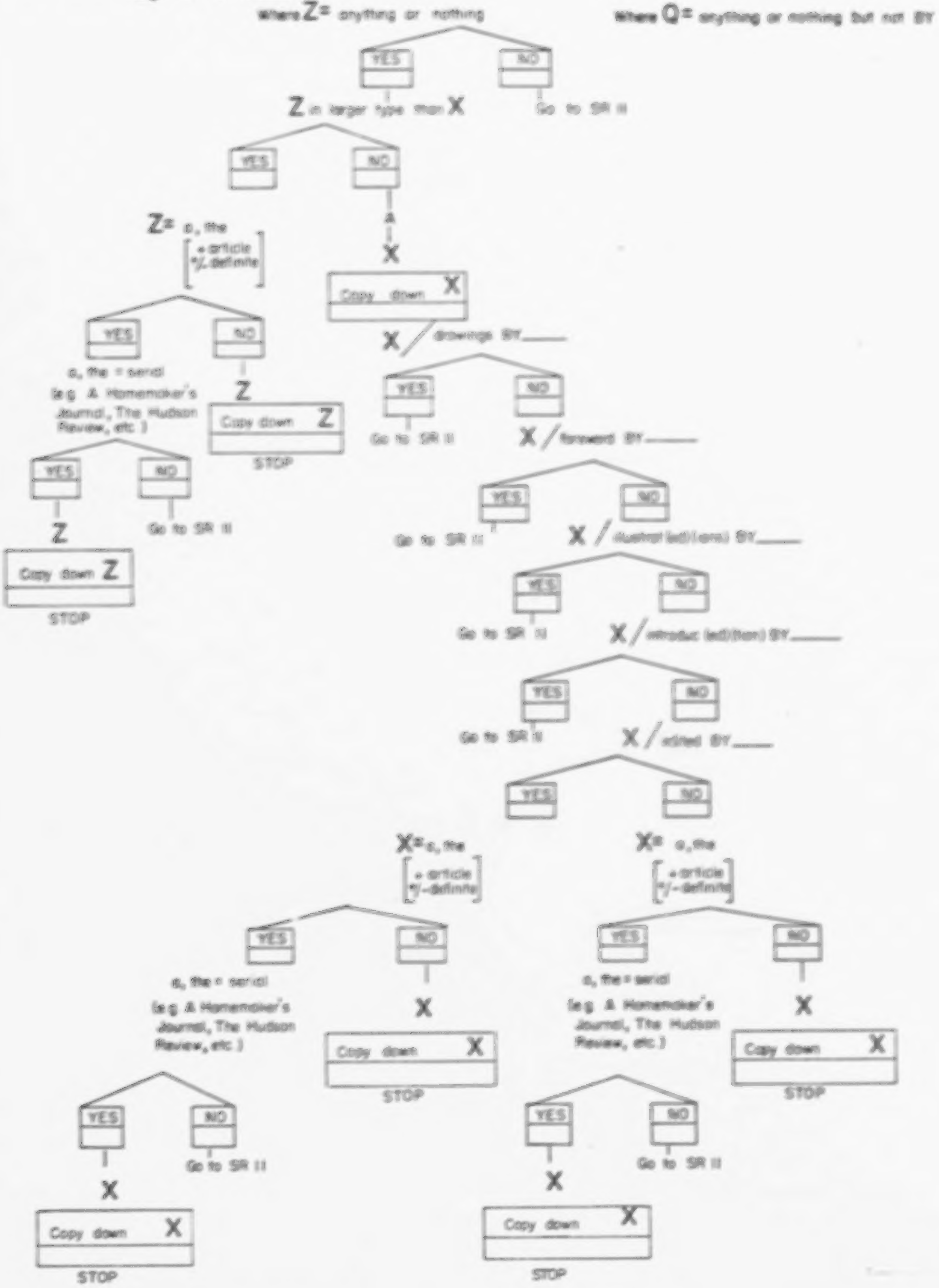


Q
X
Y
Y
Y
Z
YES
NO
Copy down
STOP


Q X Y Y Y Z X Z
yes
no
stop

ALASR 10
Recognition

Q BY XZ

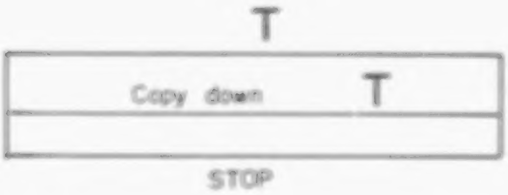

Where Z = anything or nothing
Where Q = anything or nothing but not BY
YES
NO
Go to SR II
Z in larger type than X
YES
NO
Z = a, the
[+ article
+/- definite]
a
X
Copy down X
X / drawings BY____
YES
NO
a, the = serial
(e.g. A Homemaker's Journal, The Hudson Review, etc.)
Z
Copy down Z
STOP
YES
NO
Go to SR II
X / foreword BY____
YES
NO
Z
Copy down Z
STOP
Go to SR II
YES
NO
Go to SR II
X / illustrat(ed)(ions) BY____
YES
NO
Go to SR II
X / introduc(ed)(tion) BY____
YES
NO
Go to SR II
X / edited BY____
YES
NO
X = a, the
[+ article
+/- definite]
X = a, the
[+ article
+/- definite]
YES
NO
YES
NO
a, the = serial
(e.g. A Homemaker's Journal, The Hudson Review, etc.)
X
Copy down X
STOP
a, the = serial
(e.g. A Homemaker's Journal, The Hudson Review, etc.)
X
Copy down X
STOP
YES
NO
Go to SR II
X
Copy down X
STOP
YES
NO
Go to SR II
X
Copy down X
STOP

Everything else → T

Where Everything else = recognitions to which previous rules did not apply

→ = goes to

T = Title


T
Copy down T
STOP

ALA Selection Rules

SR 1 $$X \overset{\&}{\text{AND}} Y \text{'S,S'} Z \text{ BY } W Z \longrightarrow \text{A}$$

Where $X, Y, W =$ author names

$Z =$ anything or nothing

And where A $= X \neq \text{a} \begin{bmatrix} + \text{article} \\ - \text{definite} \end{bmatrix}$ unless a serial (e.g. A Travel & Camera's etc.)

SR 2 $$X \text{'S,S'} Z \longrightarrow \text{A}$$

Where $X =$ an author name

$Z =$ anything or nothing but not BY

And where A $= X \neq \text{a} \begin{bmatrix} + \text{article} \\ - \text{definite} \end{bmatrix}$ unless a serial (e.g. A Homemaker's Journal, etc.)

SR 3 $$X \text{'S,S'} Z \begin{Bmatrix} \text{drawings} \\ \text{foreword} \\ \text{illustrat(ed)(ions)} \\ \text{introduc(ed(tion)} \\ \text{to...edited} \end{Bmatrix} Q \text{ BY } Y Q \longrightarrow \text{A}$$

Where $X, Y =$ author names

$Z =$ anything or nothing but not BY

$Q =$ anything or nothing

And where $\text{A} = X \neq \text{a} \begin{bmatrix} + \text{article} \\ - \text{definite} \end{bmatrix}$ unless a serial (e.g. A Homemaker's Journal, etc.)

$$\text{SR 4 OF} \;/\; \text{Life and} \left\{\begin{array}{l}\text{articles}\\ \text{conversations}\\ \text{correspondence}\\ \text{dialogues}\\ \text{diary, diaries}\\ \text{essays}\\ \text{journal, journals}\\ \text{letters}\\ \text{poems}\\ \text{works}\\ \text{writings}\end{array}\right\} \_\; XZ \text{ by } YZ \longrightarrow A$$

Where X, Y = author names
Z = anything or nothing

And where A = $Y \neq$ a, the $\begin{bmatrix}+\text{ article}\\ +/-\text{ definite}\end{bmatrix}$ unless a serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)

$$\text{SR 5} \left\{\begin{array}{l}\text{OF}\\ \text{FROM}\\ \text{BETWEEN}\\ \text{WITH}\end{array}\right\} \Big/ \left\{\begin{array}{l}\text{articles}\\ \text{conversations}\\ \text{correspondence}\\ \text{dialogues}\\ \text{diary, diaries}\\ \text{essays}\\ \text{journal, journals}\\ \text{letters}\\ \text{poems}\\ \text{works}\\ \text{writings}\end{array}\right\} \_\; XZ \longrightarrow A$$

Where X = author name
Z = anything or nothing

And where A = $X \neq$ a, the $\begin{bmatrix}+\text{ article}\\ +/-\text{ definite}\end{bmatrix}$ unless a serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)

SR 6 $Q_{\text{BY}}\, X_{\text{OF}}\, YZ \longrightarrow A$

Where X, Y = author names
Z = anything or nothing
Q = anything or nothing but not BY or drawings, foreword; illustrat(ed)(ions), introduc(ed)(tion)

And where $A = X \neq$ a, the $\begin{bmatrix} + \text{article} \\ +/- \text{definite} \end{bmatrix}$ unless a serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)

Otherwise Y

Where Y = a serial (e.g. Life, Playboy, Ski, etc.)

Unless Z or Y is in larger type

and $\neq$ a, the $\begin{bmatrix} + \text{article} \\ +/- \text{definite} \end{bmatrix}$ unless a serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)

SR 7 $Q_{\text{BY}}\, X_{\text{AND}}\, Y_{\text{OF}}\, WZ \longrightarrow A$

Where X, Y, W = author names
Z = anything or nothing
Q = anything or nothing but not BY or drawings, foreword, illustrat(ed)(ions), introduc(ed)(tion)

And where $A = X \neq$ a, the $\begin{bmatrix} + \text{article} \\ +/- \text{definite} \end{bmatrix}$ unless a serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)

Otherwise W

Where W = a serial (e.g. Life, Playboy, Ski, etc.)

Unless Z or Y or W is in larger type

and $\neq$ a, the $\begin{bmatrix} + \text{ article} \\ \pm \text{ definite} \end{bmatrix}$ unless a serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)

SR 8 $Q \text{ BY } X \text{ AND}_1 Y_1 \text{ AND}_2 Y_2 \text{ AND}_3 Y_3 \; Z \longrightarrow T$

Where X, Y = author names
Z = anything or nothing but not BY
Q = anything or nothing but not BY or drawings, foreword, illustrat(ed)(ions), introduc(ed)(tion)

And where T = title

Unless Z or X or Y_1, Y_2, Y_3 is in larger type

and $\neq$ a, the $\begin{bmatrix} + \text{ article} \\ \pm \text{ definite} \end{bmatrix}$ unless a serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)

SR 9 $Q \text{ BY}_1 X_1 \text{ AND}_1 Y_1 \text{ AND}_2 Y_2 \text{ AND}_3 Y_3 \text{ edited BY}_2 X_2 \; Z \longrightarrow A$

Where X, Y = author names
Z = anything or nothing including BY but not edited BY
Q = anything or nothing but not BY or drawings, edited, foreword, illustrat(ed)(ions), introduc(ed)(tion)

And where A = X /edited BY___

Where X $\neq$ a, the $\begin{bmatrix} + \text{ article} \\ \pm \text{ definite} \end{bmatrix}$ unless a serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)

Unless Z or X_1 or Y_1, Y_2, Y_3 is in larger type

and $\neq$ a, the $\begin{bmatrix} + \text{ article} \\ \pm \text{ definite} \end{bmatrix}$ unless a serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)

SR 10 $Q_{\mathrm{BY}}XZ \longrightarrow A$

With X = author name

Z = anything or nothing

Q = anything or nothing but not BY

And where $A = X$

Where $X \neq / \left\{\begin{array}{l}\text{drawings}\\ \text{foreword}\\ \text{illustrat(ed)(ions)}\\ \text{introduc(ed)(tion)}\end{array}\right\}$ BY ___

Where $X \neq$ a, the $\left[\begin{array}{l}+\ \text{article}\\ +/-\ \text{definite}\end{array}\right]$ unless a serial (e.g. A Home-maker's Journal, The Hudson Review, etc.)

Unless Z is in larger type

and $\neq$ a, the $\left[\begin{array}{l}+\ \text{article}\\ +/-\ \text{definite}\end{array}\right]$ unless a serial (e.g. A Home-maker's Journal, The Hudson Review, etc.)

SR 11 Everything else $\longrightarrow T$

Where T = title

APPENDIX C

KEY TO SYMBOLS AND TERMS USED IN THE EXPERIMENTAL RULES

To ensure maximum uniformity in the presentation of the Recognition Rules and Selection Rules developed in the study, certain notations and terminological conventions were used. Readers familiar with studies in information theory and linguistics will recognize many of these conventions. Where conventions have been borrowed from these fields, their standard use and interpretation have been preserved as much as possible. In some few cases, symbols were introduced by the investigator. For the most part, their interpretation was derived from terminology used in the AA and ALA cataloging codes.

A	author, author main entry
T	title, title main entry
W, X, Y	author names
Q, Z	anything or nothing unless otherwise defined; e.g. anything or nothing but not BY
=	is or equals; e.g. A = Z, the author is Z
≠	is not or does not equal; e.g. A if A ≠ a serial, an author main entry if the author is not a serial
/	in the environment of; e.g. OF/articles____, OF in the environment of articles
( )	optional element that may or may not be present; e.g. BY (A), by or by a
{ }	a series of environmental elements, e.g. $\left\{\begin{matrix}\text{dialogues}\\ \text{poems}\\ \text{speeches}\end{matrix}\right\}$
[ ]	list of features of an element, e.g. an article of speech, "a", with the features of article, indefinite $\begin{bmatrix}+\ \text{article}\\ -\ \text{definite}\end{bmatrix}$

→	goes to or is entered under, e.g. everything → T, everything is entered under title
#	boundary marker to indicate a physical break on the title page that comes before, between, and after the author and title statements
[BY], [AND]	recognition markers, BY, AND, implied in the author statement
initial string	first string, the title statement on the title page
sub string	second string, the author statement on the title page

APPENDIX D

INSTRUCTIONS TO PARTICIPANTS AND SAMPLE TITLE PAGE

INTRODUCTION TO THE RECOGNITION RULES

Two commonly acknowledged items of information on the title pages of books are the author and title. It is a tacit assumption that both can be easily recognized—either by position, wording or typography, or the combination of any or all of the preceding.

It is an objective of the present study to demonstrate that one of these items, namely the author, can be recognized in a completely objective mechanical manner without the use of human intuition or foreknowledge. From a practical standpoint, objective mechanical author recognition would greatly assist libraries and similar document gathering units in expediting their catalog or file entries and their acquisitions for reader use.

The method adopted to accomplish this objective of the study involves the application of a set of rules called in terms of their function Recognition Rules.

Briefly, the Recognition Rules consist of a set of so-called markers: BY, AND, WITH, OF, BETWEEN, FROM, 'S,S', that when displayed singly or in combination on the title pages of books function as signalling devices for the presence of an author name or names, as the case may be. The illustrations below indicate their general display and author recognition function on the title page.

(1) MAGAZINE DESIGN

BY

Ruairi McLean

Marker

(2) UNITED STATES NAVY AIRCRAFT

SINCE 1911

BY

Marker

Gordon Swanborough and Peter M. Bowers

Marker

(3) YOUNG FILMMAKERS

BY

Marker

Rodger Larson

with Ellen Meade

Marker

(4) CHANTICLEER

THE POEMS OF TERRY WISE

Marker

(5) THE CORRESPONDENCE

BETWEEN BENJAMIN HARRISON AND JAMES G. BLAINE

Marker Marker

(6) LETTERS FROM WILLIAM COBBETT TO EDWARD THORNTON

Marker

(7) CARDINAL NEWMAN'S

BEST PLAIN SERMONS

Marker

Since the markers can function in numerous contexts in addition to that of author name recognition, the rules place restrictions on their recognition by specifying the position and the environment in which they must occur on the title page.

With respect to position, the rules take into account the physical separation of the author and title statements and the author recognition function of the markers in these statements. With respect to environment, the rules take into consideration certain words or elements in whose context the markers are found to function in the role of author indicating devices.

A more detailed explanation of these points is contained in the following section which presents the rules and instruction in their application to the title pages selected for the experiment.

INSTRUCTIONS IN THE APPLICATION OF THE RECOGNITION RULES

Attached are two sets of Recognition Rules, each of which is to be applied to the accompanying set of title pages reproduced from English language books. Both sets of rules have the same function, although not all the same rules. The function of the rules is to recognize the presence of an author name or names on a title page of an English language book.

The explanation and instructions that follow are applicable to both sets of rules and should be studied before applying the rules to the sets of title pages.

1. The author and title statements on a title page are viewed as a sequence of two physical strings, of which the title statement is identified by the rules as the initial or first string, and the author statement is the sub or second string. Each string is composed of as many lines as there are in each statement. For example:

(1) Initial string: Introduction to Geometry
By
Substring: G. A. Dickinson

(2) Initial string: GOLF GUIDEBOOK
Basic Form and Playing Techniques
for young people

Substring: BY
GAY BREWER
with Roger Ganen

2. The rules consider each string to be in the form: # S #, where the ends of each string are to be marked with a boundary marker, #. The boundary marker represents the physical breaks on the title page that come before, between and after the author and title statements, and are a technique used by the rules to set apart and at the same time define the two strings. Taking the examples above, boundary markers would be inserted at the ends of each string and define the string either as initial (first) string or sub (second) string:

(1) # Introduction to Geometry # } initial string

By
G. A. Dickinson # } substring

(2) # GOLF GUIDEBOOK
Basic Form and Playing Technique
for young people # } initial string

BY
Gay Brewer
with Roger Ganen # } substring

3. Application of the Recognition Rules is dependent on the abstraction of the title page as noted above.

4. The Recognition Rules are a set of instructions of the form:

Recognize such and such an element . . .

Where to "recognize" is to identify the exact shape and form of said element in the initial string, substring, or both strings as dictated, and where the element is any one of a set of markers (BY, AND, WITH, OF, BETWEEN, FROM, 'S,S') given in the rules. A simple recognition rule would be:

Recognize WITH in a substring

In applying the rule to the substring derived in example (2), the following recognition would take place:

with Roger Ganen

5. Where marker recognition is restricted to certain contexts only, the rule is of the form:

Recognize such and such an element
when it occurs in the environment of . . .

where "in the environment of" is formally expressed in the rules as a slash (/) of the type:

$$X/Y\ldots\underline{\qquad}$$

with X and Y being given marker elements and the rule reading:

Recognize marker X in the environment of marker Y, the dots, ..., representing anything or nothing coming after Y, and the blank, ___, in the rule specifying the position of marker X in the environment. An example of a context-sensitive or —restricted rule might be:

Recognize OF in the substring when OF/WITH... ___ where the marker OF is recognized only when it comes after the marker WITH; the ___ indicating the position of the marker OF in the environment. In applying the rule to the substring derived below:

#   By Chester M. Zwijewski
With the assistance of June L. Fletcher #

the following recognition would take place:

# With   ...   of June L. Fletcher #

6. Often a number of environments apply to a given marker. In such cases, the environments are listed vertically within braces, { }; the braces being used to collapse what would be several rules into a single rule. For example:

Recognize OF in the initial string when

$$\text{OF} / \left\{ \begin{array}{l} \text{dialogues} \\ \text{poems} \\ \text{works} \end{array} \right\} \_\_\_$$

where the marker OF is recognized when it comes directly after "dialogues", when it comes directly after "poems", when it comes directly after "works"; its position indicated with the ___. This particular rule would apply to the strings derived below:

(1) # WORKS OF ROBERT FROST #

# Edited by Walter M. Lewis #

(2) # Selected Poems of Albert Steffen #

# Translated from German
By Daisy Alden #

to produce the respective recognitions:

(1) # WORKS OF ROBERT FROST #

(2) # Poems of Albert Steffens #

7. In addition to a vertical listing within braces, some environmental elements are enclosed in parentheses, ( ). In such cases, parentheses are used to indicate optional elements, elements that may or may not be present in the string. An example of an optional element is the indefinite article, A, which occurs in the following rule for the marker 'S, S':

> Recognize (A) ___ 'S,S' when it occurs without BY in the substring

which reads: recognize the marker 'S,S' by itself or with the association of the indefinite article, A, when present. Application of the rule to the strings derived below:

(3) # A Pilgrim's Journey to Nirvana #

(4) # Nairn's London #

would produce the respective recognitions:

(3) # A Pilgrim's #

(4) # Nairn's #

8. With two exceptions, the markers and environments specified in the rules must have identical physical expression in the strings for recognition of the author name to take place. The exceptions occur with the markers BY and AND, which in many cases are implied although not physically expressed in the author statement. The rules for these markers take cognizance of this condition and are stated in the form:

> Recognize BY, as implied, in the substring, etc.

In the case of BY, the marker is said to be implied when a substring is present and there is no physical expression of the marker in it. For example:

(5) THE HIDDEN DIMENSION

Edward T. Hall

BY implied

In such cases, the rule dictates that the marker be inserted in brackets, [ ], at the head of the substring. Using the illustration above, the rule would apply to produce the following recognition:

[BY] Edward T. Hall

In general, the marker AND is said to be implied when there is a substring and no physical expression of this or any other marker except BY at the head of the substring. For example:

(6) Underground Storage of Fluids

By
Donald L. Katz

Keith H. Coats

AND implied

In such cases, the rule dictates insertion of the marker at the end of each string except the final string in the substring. Application of the rule to the illustration above would produce the recognition:

By Donald L. Katz [AND] Keith H. Coats

In some title page displays, the marker AND is implied in the form of a slash, /, or a comma. For example:

(7) Introduction to Moral Education

By
John Wilson / Norman Williams / Barry Sugarman

AND implied

(8) How To Grow Plants – Indoors and Outdoors

By
Gerald Whitney, James Gooch

and Doris Royce

AND implied

In these cases, the rule dictates replacement of the symbols with the insertion of the marker word. When applied to the illustrations above, the rule would produce the respective recognitions:

(7) # By John Wilson [AND] Norman Williams [AND]
Barry Sugarman #

(8) # By Gerald Whitney [AND] James Gooch and
Doris Royce #

8. Finally, with respect to their application, the rules are strictly ordered. Application starts at rule 1 and continues with the application of successive rules 2, 3, and so on through to the last rule, this procedure being followed for each title page. With the application of each rule, the appropriate strings are examined for the marker and environments in question. If the marker and environments are identical to those specified in the rule, the recognition is copied down. With the application of each succeeding rule, the recognitions copied down from the previous rules are carried forward and noted in the appropriate strings and in the order in which they appear in their strings. Two lines for the two strings (initial and sub) are provided under each rule for this purpose. For example, given the title page display:

Letters from WILLIAM COBBETT to EDWARD THORNTON

Edited By

G. D. H. Cole

under Recognition Rule 1, BY in the substring would be recognized, copied down, and set off with appropriate boundary markers in the line provided for the substring:

____

__#______By G. D. H. Cole__________#____

With the application of Recognition Rule 4, FROM would be recognized since it is in the environment of LETTERS ____. This recognition would be copied down and set off with the appropriate boundary markers in the initial string:

__#__letters from WILLIAM COBBETT____#____

__#__By G. D. H. Cole_______#__________

For examples of the application of the Recognition Rules to sample title page displays, see the following:

Facsimile Title Page:

FARM AND PERSONAL FINANCE
ESSAYS

By
John J. Pianko

Application of Recognition Rule 1 (a) in both sets of rules would produce the following recognition, copied down, and set off with appropriate boundary markers in the appropriate string:

#______________________________#

#____By John J. Pianko__________#

With the application of the remaining rules, this same recognition would be carried forward.

Facsimile Title Page:

Heroes and Heroines

By
Emma Tully
with Hazel Stone

Application of Recognition Rule 1 (a) in both sets of rules would produce the following recognition, copied down, and set off with appropriate boundary markers in the appropriate string:

#__________________________#

#_____By Emma Tully__________#

With the application of Recognition Rule 7 in both sets of rules, the recognition would be completed:

#______________________________#

#_By Emma Tully with Hazel Stone_#

INTRODUCTION TO THE SELECTION RULES

One of the chief functions of a library catalog is to record each work by author. Provision of the entry presumes the recognition and selection of the person or body responsible for a work.

It is an objective of the present study to demonstrate that the recognition of an author and the identification of author responsibility for a work can be accomplished in a completely objective mechanical manner without the use of human intuition or foreknowledge. From a practical standpoint, author recognition and selection would greatly assist libraries in preparing such entries for their catalogs and in expediting their acquisitions for reader use.

The method adopted to accomplish this objective of the study involves the application of two sets of rules, called in terms of their function, Recognition Rules and Selection Rules. A description of the Recognition Rules and instructions in their application were presented in the previous section. In the present section, the Selection Rules will be presented followed by a listing of the rules and instructions in their practical application.

The Selection Rules in question consist of a set of instructions for determining author responsibility. On the basis of the author name recognition produced under the application of the Recognition Rules, the Selection Rules operate to chose the appropriate entry, either an author entry or title entry.

With respect to the author entry, the principle underlying its use is that one person is assigned the entry and that that person is considered to be chiefly responsible for the work. In the absence of authorship or authorship responsibility, entry is under the title of the work.

On the following pages, the rules, instructions and illustrations in their practical application are presented in detail.

INSTRUCTIONS IN THE APPLICATION OF THE SELECTION RULES

Attached are two sets of Selection Rules, each of which is to be applied to the appropriate set of title pages to which the Recognition Rules have been applied. Both sets of rules have the same function, although not all the same rules. The function of the rules is to select the main entry or heading for a work.

The explanation and instructions that follow are applicable to both sets of rules and should be studied before applying the rules to the sets of title pages indicated.

1. The Selection Rules apply after the Recognition Rules and are of the form:

$$X \longrightarrow A$$

which is to be interpreted as an instruction to enter X under A. Given the markers and author names copied down by the Recognition Rules, entry selection is specified as A (author) or T (title). For example, given the recognition:

$$\text{BY } X \text{ AND } Y$$

where X and Y represent author names copied down by the Recognition Rules, entry selection is specified as:

$$\text{BY } X \text{ AND } Y \rightarrow A$$

$$\text{Where } A = X$$

and reads: where there is marker BY X (any author name), marker AND Y (any author name), the entry is A (author), where the author is X .

2. Since there may be a multiplicity of markers in any given recognition, the selection rules utilize two symbols: Q and Z to represent these complexities. These symbols are incorporated in the rules in the form:

$$\text{BY } X Z \longrightarrow A$$

Where $A = X$
And Z = Anything or nothing

with the interpretation that where there is marker BY X (any author name) by itself or with Z, any other author name or names in a recognition, entry is under A (author), where the author is X. This rule would be applicable to recognitions such as:

By John Smith and Mary Smith

where X is John Smith and Z is Mary Smith, in which case entry is under X, John Smith.

3. Where an entry applies only when certain conditions are met, the rule is in the form:

$$X / \text{drawings BY } \_\_\_ \longrightarrow T$$

Where T = Title of monograph

where the slash in the rule, $/$, indicates the environment in which X (any author name) must occur, and the blank, ___, indicates the position of the author name in the environment. Given the recognition:

# Drawings by Lettie Wolf #

this particular rule would apply to assign a title entry.

4. Often the same entry is designated for a number of environments. In such cases, the environments are listed vertically within braces, $\{\ \}$. As in the case of the recognition rules, the braces are used to collapse what would be several rules into a single rule. For example:

$$\text{OF } / \left\{ \begin{array}{l} \text{dialogues} \\ \text{poems} \\ \text{works} \end{array} \right\} \_\_\_ X Z \longrightarrow A$$

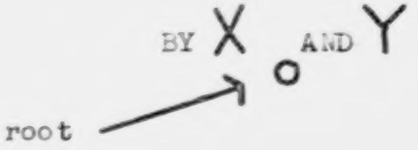
Where $A = X$
And Z = anything or nothing

which reads: where the marker OF X (any author name) by itself or with any other author name comes after "dialogues" or "poems" or "works", the entry is A (author) where A is X. With a recognition such as:

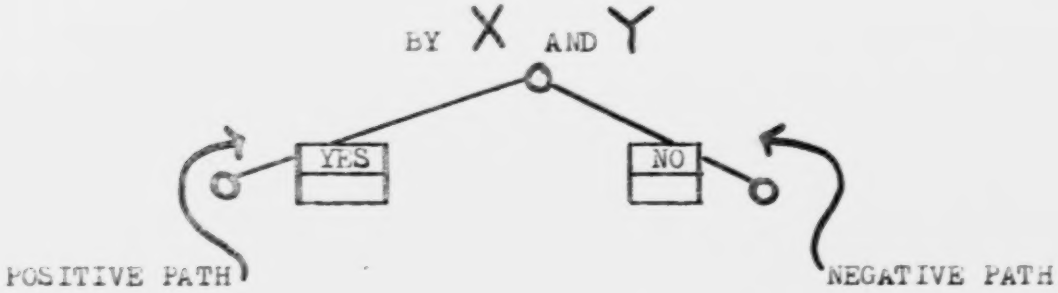
WORKS OF ROBERT FROST

this rule would apply to assign entry under Robert Frost.

5. To enable the selection rules to accept a recognition and to examine it for certain physical characteristics, the rules are represented graphically in the form of "decision trees". At the root or highest node of each tree is the construct of the rule. For example:

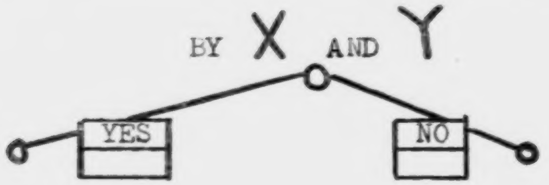


From the root, two lines or branches representing positive and negative paths extend downward to two lower nodes



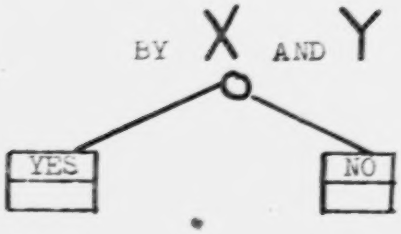
so that if a recognition produced under the Recognition Rules should match the construct of the selection rule at the root, the recognition would be passed to the node to the lower left. If the recognition should not match the construct at the root, the recognition would be passed to the node to the lower right. An identical match is indicated below for the recognition:

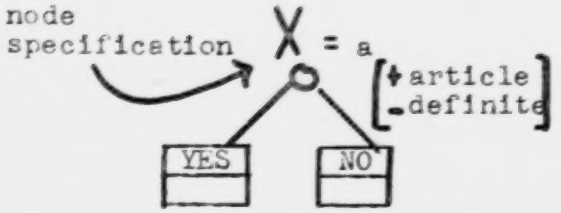
# By John Smith and Mary Smith #



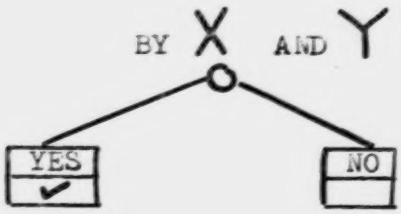
where X equals John Smith and Y equals Mary Smith.

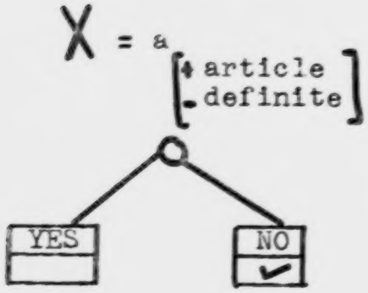
Associated with each node is a specification that the recognition must satisfy positively or negatively


BY X AND Y
YES
NO

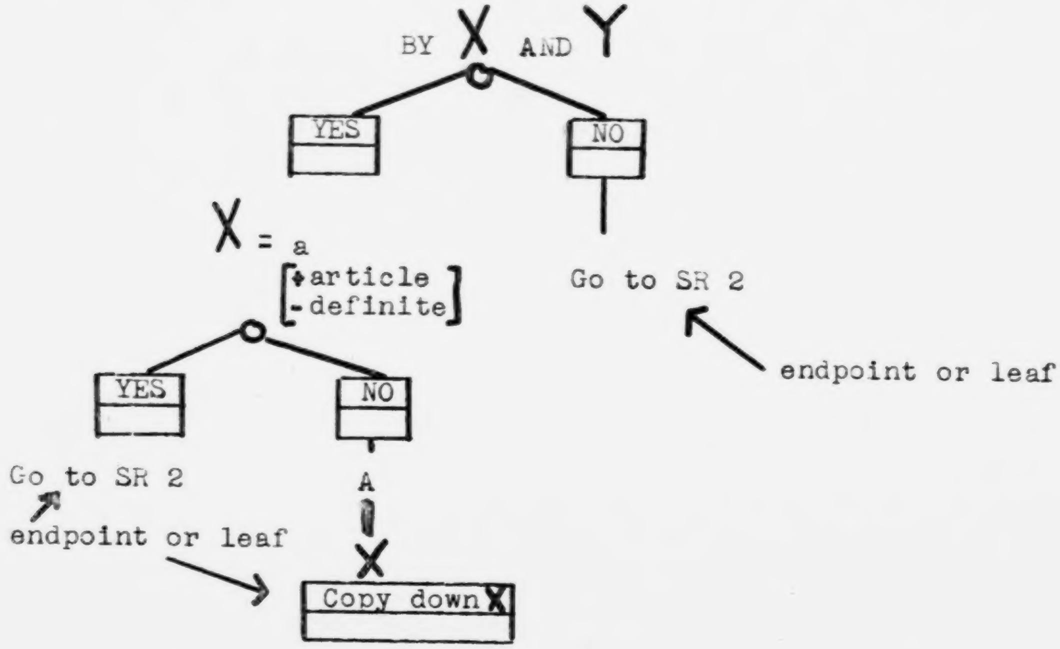

node specification
X = a
+ article
- definite
YES
NO

With each node, the recognition produced under the Recognition Rules must be examined for the specification. For the recognition: # By John Smith and Mary Smith #, the specification is indicated as negative; that is, X (John Smith) in the recognition is not introduced with the indefinite article, a:

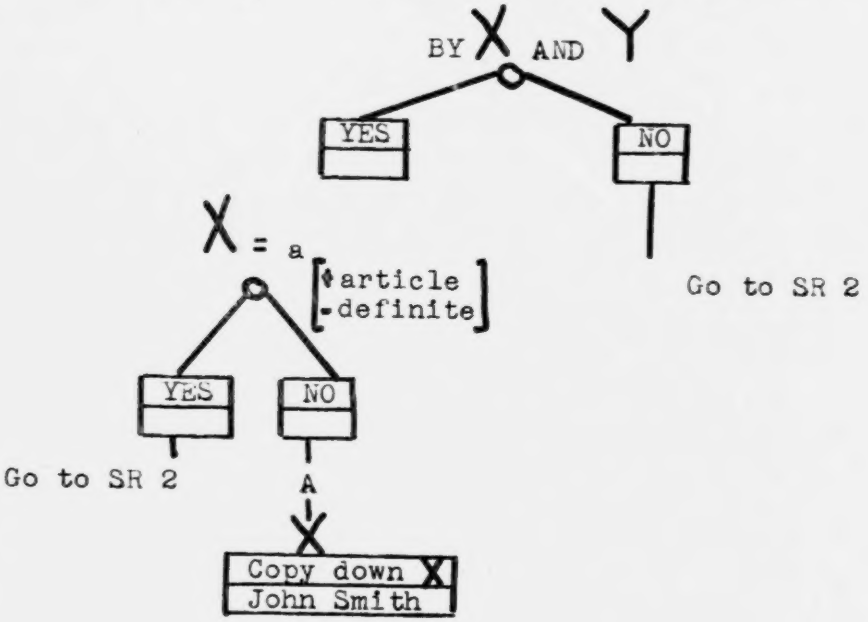

BY X AND Y
YES
✓
NO


X = a
+ article
- definite
YES
NO
✓

Entry selection continues in this manner until an endpoint or leaf is reached. This endpoint or leaf is a node with no nodes below it. The endpoints of the Selection Rules dictate the entry choice for the recognition, A(author) or T(title) or direct the recognition to the next rule.


BY X AND Y
YES
NO
Go to SR 2
endpoint or leaf
X = a
[+article -definite]
YES
NO
Go to SR 2
endpoint or leaf
A
X
Copy down X

In the case of the earlier recognition: # By John Smith and Mary Smith #, the endpoint in the Selection Rule above would dictate an A(author) entry under John Smith:


BY X AND Y
YES
NO
Go to SR 2
X = a
[+article -definite]
YES
NO
Go to SR 2
A
X
Copy down X
John Smith

However, if the recognition did not match the construct of the selection rule at the root, or did not satisfy the specifications at the nodes, the endpoints of the rule would require that the recognition in question be examined by the next rule.

6. Finally, with respect to their application, the selection rules are arranged to be applied to a recognition in a given order: 1, 2, 3... Application starts at Rule 1 and continues with the application of successive rules only until the specifications of entry are satisfied. As soon as an endpoint with an entry designation is reached, the selection process terminates. In the trees, these endpoints are marked with the word STOP, indicating termination of the process. For an example of the application of Selection Rule 1 to a recognition produced under the recognition rules, see the following page.

EXAMPLE

Facsimile Title Page:

Pratt and MacKenzie's

LAW OF HIGHWAYS

21st Edition
By Harold Parrish and Lord de Mauley

Application of Recognition Rule 2(a) in both sets of rules produces the following recognition, copied down and set off with the appropriate boundary markers in the appropriate strings:

# Pratt and Mackenzie's #

# By Harold Parrish #

Application of Recognition Rule 6(c) in the respective sets of rules would produce recognition of the marker AND in the substring, and this recognition would be copied down and the recognition for the previous rule above would be carried forward, set off with appropriate boundary markers in the appropriate strings:

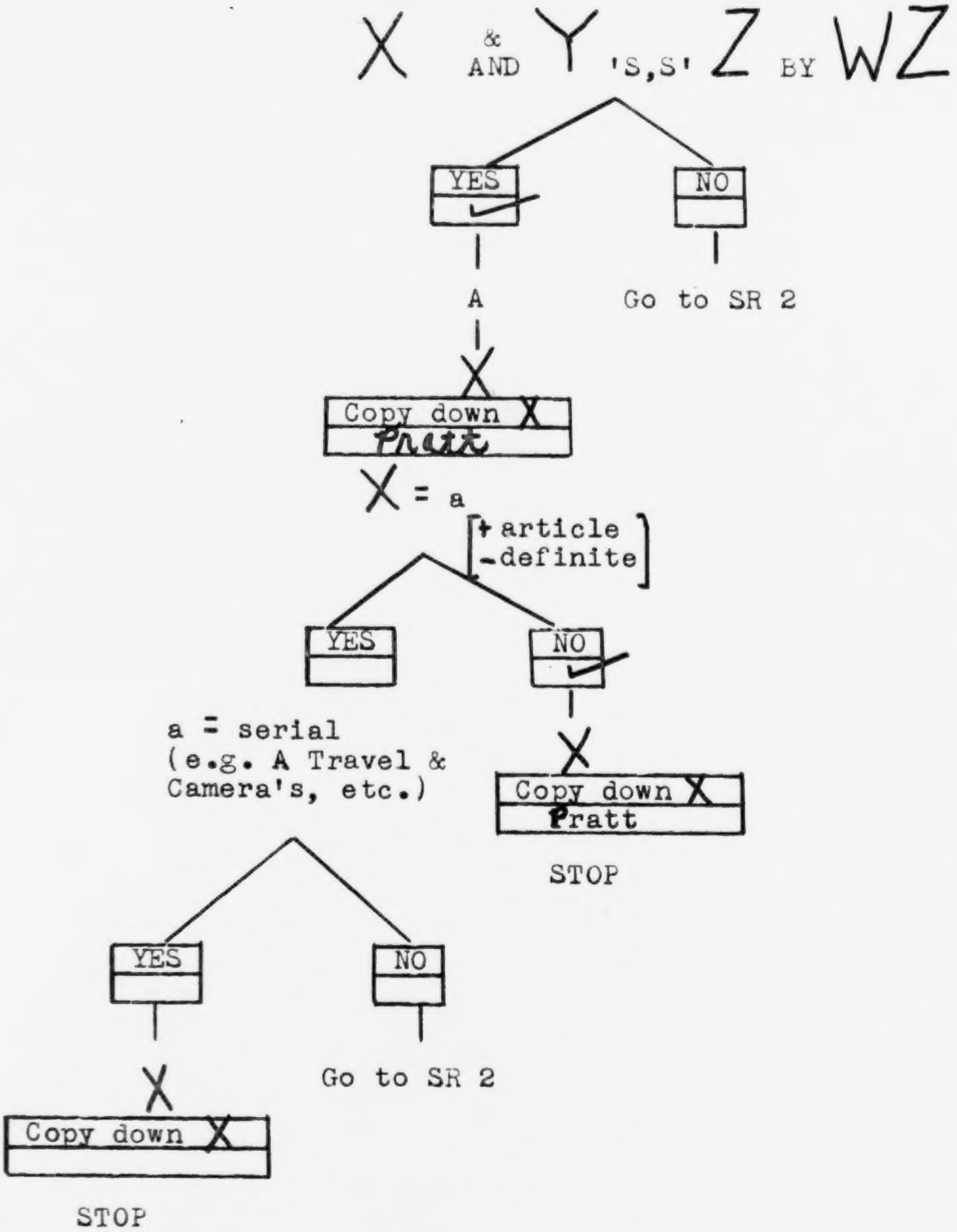
# Pratt and MacKenzie's #

# By Harold Parrish and Lord de Mauley #

After the Recognition Rules, the Selection Rules would be applied, starting with S(election) R(ule) 1. In our example, the last recognition copied down under the Recognition Rules would be examined to see if it matches the construct of the selection rule at the root of the tree. This examination requires that you treat the initial and substrings in the recognition as one continuous string:

Pratt and MacKenzie's By Harold Parrish and
Lord de Mauley

Application of ALA SR 1 to the recognition produces the following entry selection for the work:


X & AND Y 'S,S' Z BY WZ
YES
✓
NO
A
Go to SR 2
X
Copy down X
Pratt
X = a
+ article
- definite
YES
NO
✓
a = serial
(e.g. A Travel & Camera's, etc.)
X
Copy down X
Pratt
STOP
YES
NO
X
Go to SR 2
Copy down X
STOP

In this particular example, examination of the recognition with the construct of the selection rule indicates a match, in which case YES is checked.

The entry is designated as A(author), where the author is X or Pratt in the recognition. Pratt is copied down and next examined for the specification at the node.

The specification is indicated as negative, since Pratt is not introduced with the indefinite article, a.

Satisfying the node specification, entry is designated under Pratt, and the selection process for this recognition terminates.

Sample Title Page

Start → *Letters from the Saints*

arranged and selected by a
Benedictine of Stanbrook Abbey
with illustrations by Vivian Berger ← Stop

[illegible] BOOKS, INC. PUBLISHERS NEW YORK

ALA Recognition Rules

RR 1(a) Recognize BY in any substring

Recognition [✓ YES] NO

If YES, copy down recognition for the rule, set off with appropriate boundary markers in the appropriate string

____

by a Benedictine by Vivian Berger

Go to RR 1(b)

If NO, go to RR 1(b)

RR 1(b) Recognize BY as implied at the head of any substring in which there is no physical expression of the marker in it

Recognition YES [✓ NO]

If YES, copy down recognition for the rule with BY in brackets, [ ], set off with appropriate boundary markers in the appropriate string

____

____

Go to RR 1(c)

If NO, go to RR 1(c)

RR 1(c) Recognize

$$\left\{\begin{array}{l}\text{drawings}\\ \text{edited}\\ \text{foreword}\\ \text{illustrat(ed)(ions)}\\ \text{introduc(ed)(tion)}\end{array}\right\} \Big/ \_\_\_\_ \text{ BY}$$

In any substring

Recognition [✓ YES] NO

If [YES], copy down recognition for the rule and carry recognition for the previous rule forward, set off with appropriate boundary markers in the appropriate strings

____

#by a Benedictine; illustrations by Vivian Berger#

Go to RR 2(a)

If [NO], go to RR 2(a)

RR 2(a) Recognize (A) ___$\overset{\&}{\text{AND}}$___'S,S' when

(A)___$\overset{\&}{\text{AND}}$___'S,S' / ___...BY [BY] ...

Recognition YES ☐ NO ☑

If [YES], copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 2(b)

If [NO], go to RR 2(b)

RR 2(b) Recognize (A)___'S,S' when

(A)___'S,S' /___... { drawings
foreword
illustrat(ed)(ions)
introduc(ed)(tion)
to...edited } ...BY [BY]...

Recognition YES ☐ NO ☑

If YES, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 2(c)

If NO, go to RR 2(c)

RR 2(c) Recognize (A)___'S,s' when it occurs without BY [BY] in the substring

Recognition YES ☐ NO ☑

If YES, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 3

If NO, go to RR 3

RR 3 Recognize OP when

$$\text{OP} \;/\; \text{Life and} \left\{\begin{array}{l}\text{articles}\\ \text{conversations}\\ \text{correspondence}\\ \text{dialogues}\\ \text{diary, diaries}\\ \text{essays}\\ \text{journal, journals}\\ \text{letters}\\ \text{poems}\\ \text{screenplays}\\ \text{works}\\ \text{writings}\end{array}\right\} \_\_\_ \ldots$$

Recognition YES ☐ NO ☑

If YES, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 4

If NO, go to RR 4

RR 4 Recognize OF, FROM, BETWEEN, WITH when

$$\left\{\begin{array}{l}\text{OF}\\ \text{FROM}\\ \text{BETWEEN}\\ \text{WITH}\end{array}\right\} \Big/ \left\{\begin{array}{l}\text{articles}\\ \text{conversations}\\ \text{correspondence}\\ \text{diary, diaries}\\ \text{essays}\\ \text{journal, journals}\\ \text{letters}\\ \text{poems}\\ \text{screen plays}\\ \text{works}\\ \text{writings}\end{array}\right\} \_\_\_ \ldots$$

Recognition YES ☑ NO ☐

If YES, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

Letters from the Saints

by a Benedictine illustrations by Vivian Berger

Go to RR 5

If NO, go to RR 5

RR 5 Recognize OF after BY in the substring

Recognition ☑ YES ☐ NO

If YES, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

# Letters from the Saints #

# by a Benedictine of Stanbrook Abbey illustrations by Vivian Berger #

Go to RR 6(a)

If NO, go to RR 6(a)

RR 6(a) Recognize AND as implied at the end of each string except the final string in the substring when:

(a) BY [BY] is at the head of the substring and there is no other physical expression of the marker BY or the markers AND, OF, or WITH in the substring

(b) BY [BY] is at the head of the substring and EDITED BY is at the end of the substring and there is no other physical expression of the marker BY or the markers AND, OF, or WITH in the substring

Recognition ☐ YES ☑ NO

If YES, copy down recognition for the rule with AND in brackets, [ ], and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

Go to RR 6(b)

If NO, go to RR 6(b)

RR 6(b) Recognize AND, as implied in the form of a comma (,) or slash (/) in the substring, when:

(a) BY [BY] is at the head of the substring and there is no other physical expression of the marker BY or the markers OF or WITH in the substring

(b) BY [BY] is at the head of the substring and EDITED BY is at the end of the substring and there is no other physical expression of the marker BY or the markers OF or WITH in the substring

	YES	NO
Recognition		✓

If $\boxed{\text{YES}}$, copy down recognition for the rule with AND in brackets,   , and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 6(c)

If $\boxed{\text{NO}}$, go to RR 6(c)

RR 6(c) Recognize AND, &, after BY in the substring, except when:

(a) $\left\{\begin{matrix}\text{AND}\\ \&\end{matrix}\right\} \Big/ \text{BY}[\text{BY}]\ldots\text{WITH}\ldots\_\_$

(b) $\left\{\begin{matrix}\text{AND}\\ \&\\ \text{AND}\end{matrix}\right\} \Big/ \text{BY}[\text{BY}]\ldots\left\{\begin{matrix}\text{edited}\_\_\\ \_\_\text{edited}\end{matrix}\right\}\ldots$

	YES	NO
Recognition		✓

If $\boxed{\text{YES}}$, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

Go to RR 7

If NO, go to RR 7

RR 7 Recognize WITH after BY in the substring except when

(a) WITH / BY [BY]...___...BY [BY]

(b) WITH / BY [BY]...___...OF

Recognition YES ☐ NO ☑

If YES, copy down recognition for the rule and carry recognition for the previous rules forward, set off with appropriate boundary markers in the appropriate strings

____

____

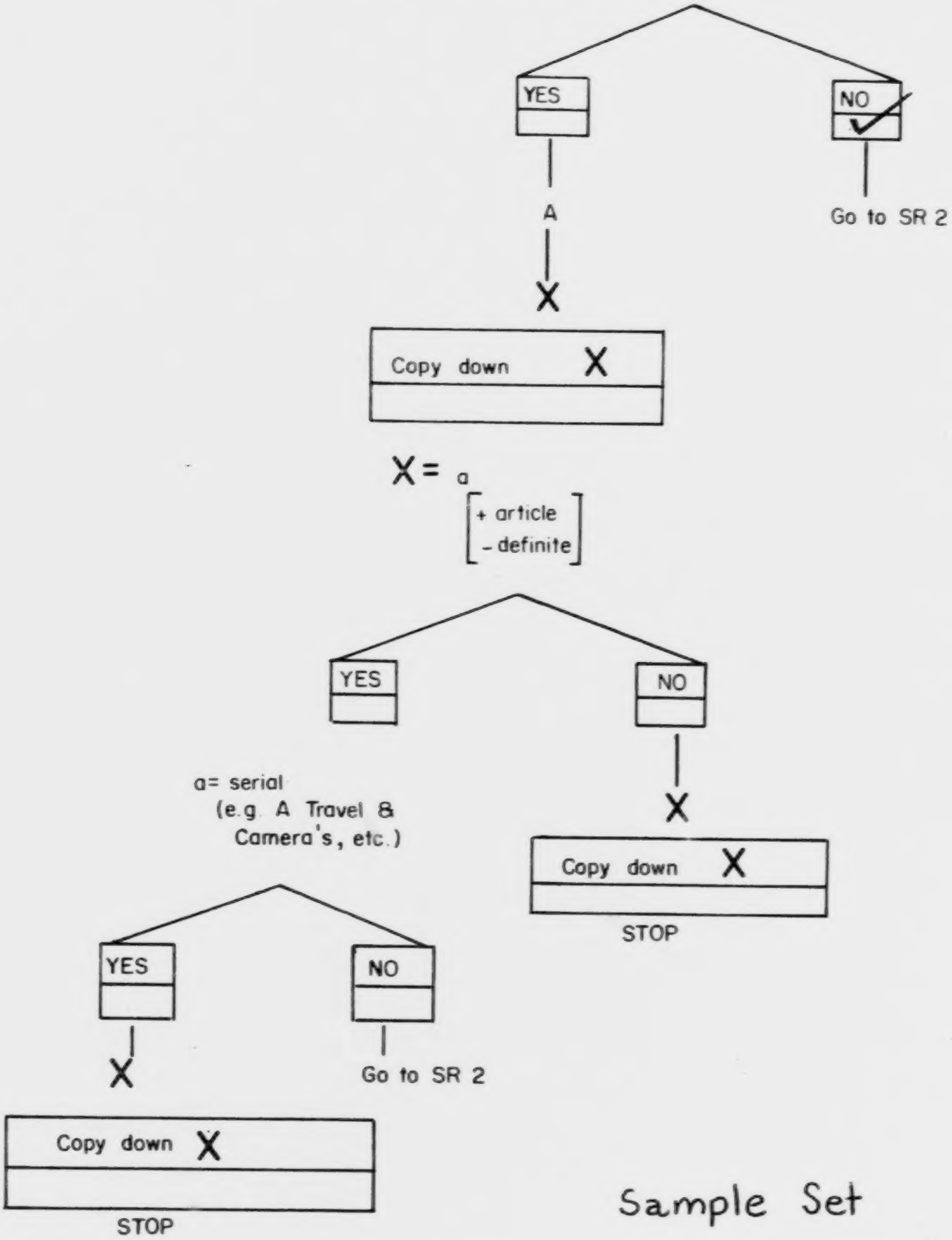
Apply ALA Selection Rules to the recognition

If NO, apply ALA Selection Rules to the last recognition copied down under the rules above or to the "non" or no recognition as the case may be

ALASR I
Recognition

X & AND Y 's, s' Z BY WZ

Where Z = anything or nothing

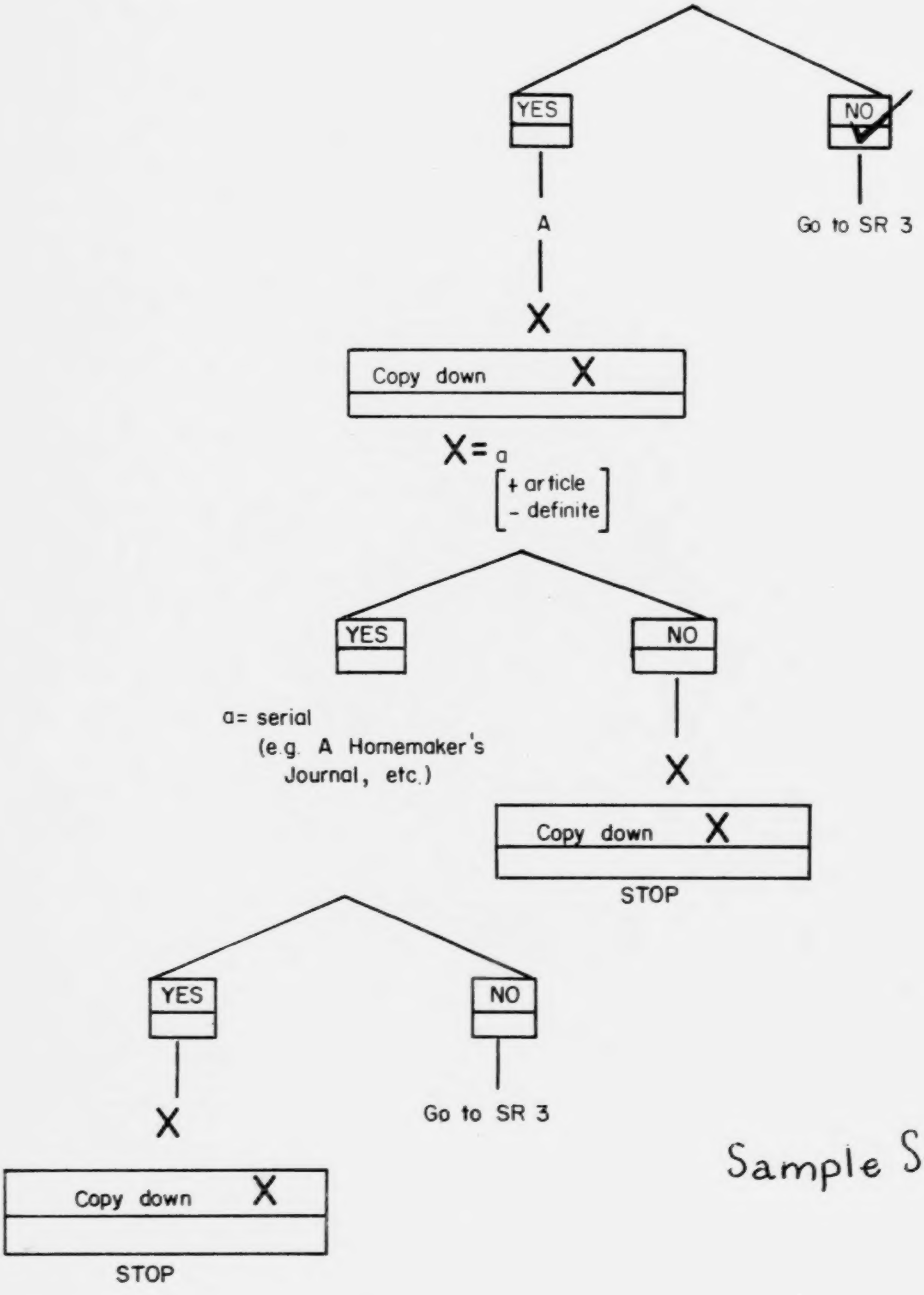

YES
NO
A
X
Go to SR 2
Copy down X
X = a
+ article
- definite
YES
NO
X
Copy down X
STOP
a= serial
(e.g. A Travel & Camera's, etc.)
YES
NO
X
Go to SR 2
Copy down X
STOP

Sample Set

ALASR 2
Recognition:

X 's, s' Z

Where Z = anything or nothing but not BY


YES
NO ✓
A
Go to SR 3
X
Copy down X
X = a
+ article
- definite
YES
NO
a= serial
(e.g. A Homemaker's
Journal, etc.)
X
Copy down X
STOP
YES
NO
X
Go to SR 3
Copy down X
STOP

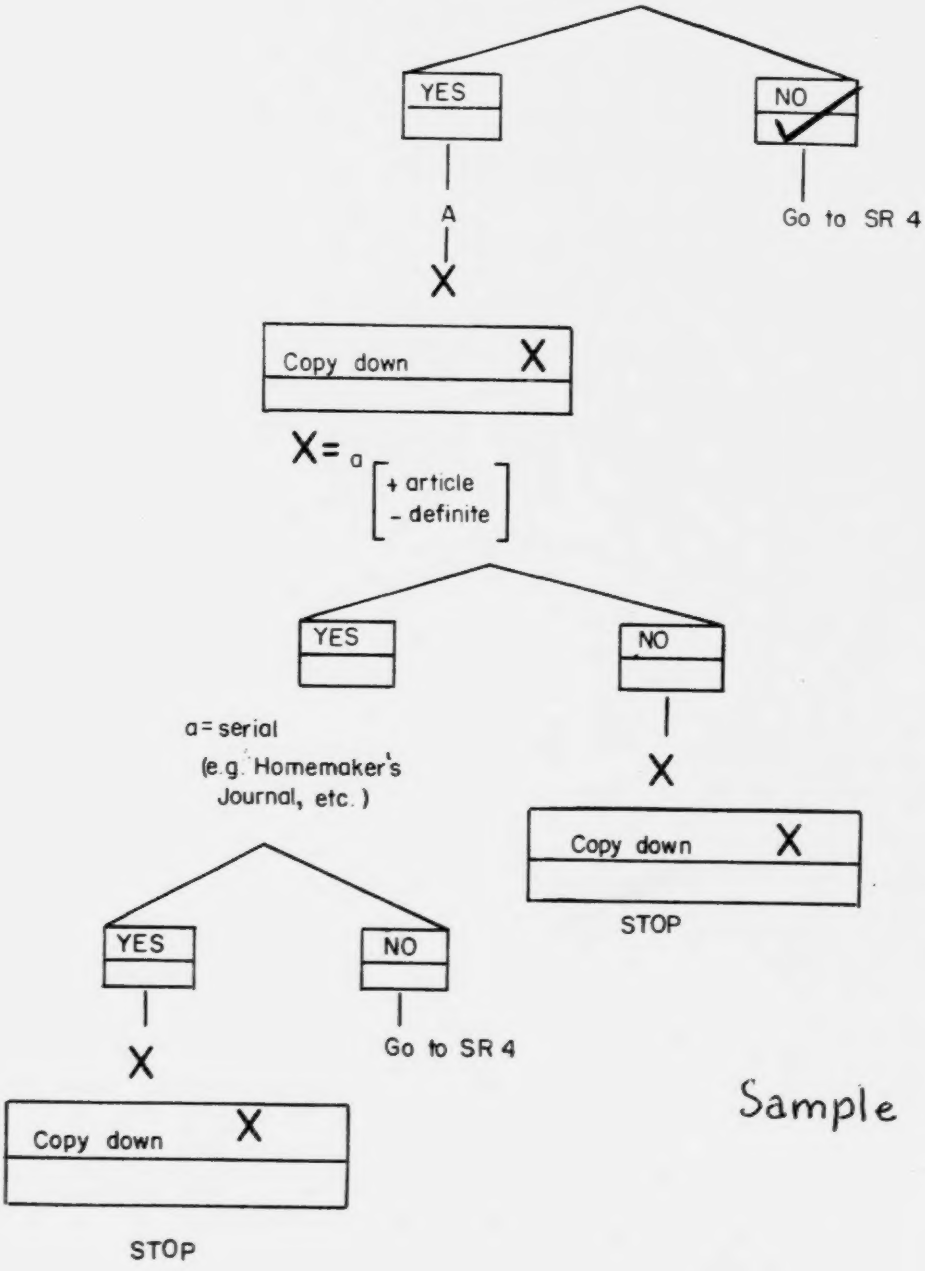
Sample Set

ALASR 3
Recognition:

$$X \text{ 's, s' } Z \left\{\begin{bmatrix}\text{drawings}\\ \text{foreword}\\ \text{illustrat (ed) (ions)}\\ \text{introduc (ed) (tion)}\\ \text{to... edited}\end{bmatrix}\right\} Q \text{ BY } YQ$$

Where Z= anything or nothing but not BY

Q= anything or nothing

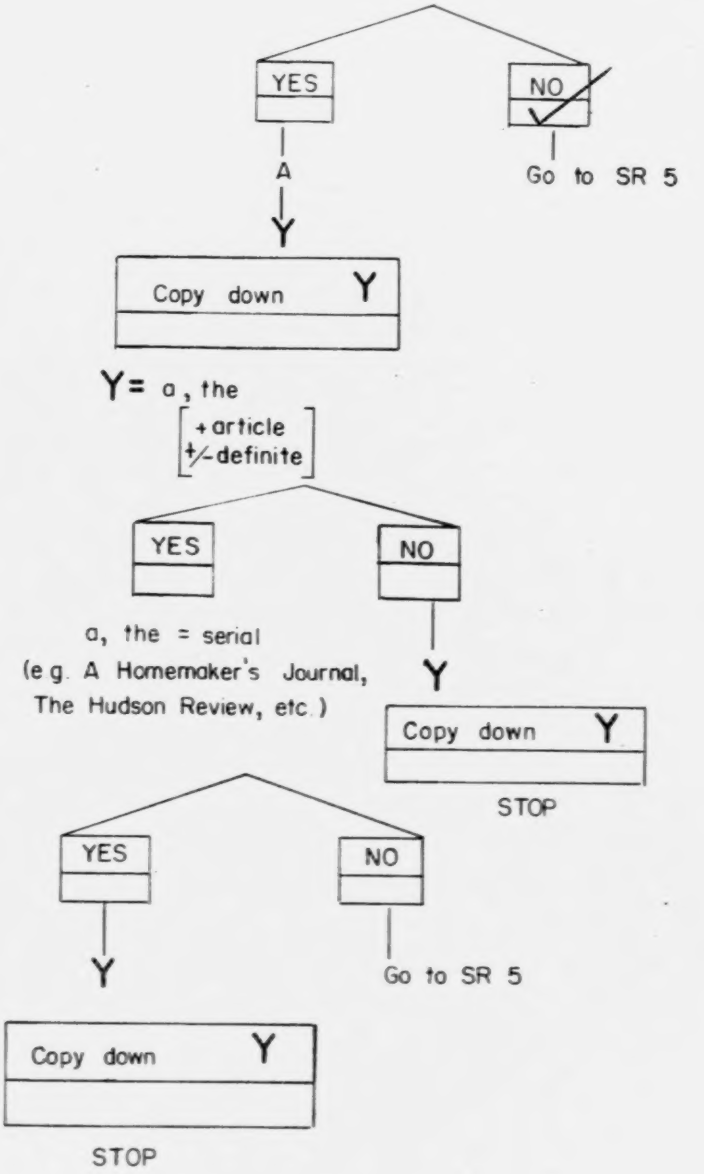

YES
NO ✓
A
X
Go to SR 4
Copy down X
X= a [+ article - definite]
YES
NO
a= serial (e.g. Homemaker's Journal, etc.)
X
Copy down X
STOP
YES
NO
X
Go to SR 4
Copy down X
STOP

Sample Set

ALASR 4
Recognition

OF / Life and {articles, conversations, correspondence, dialogues, diary, diaries, essays, journal, journals, letters, poems, screenplays, [illegible]} — XZ BY YZ

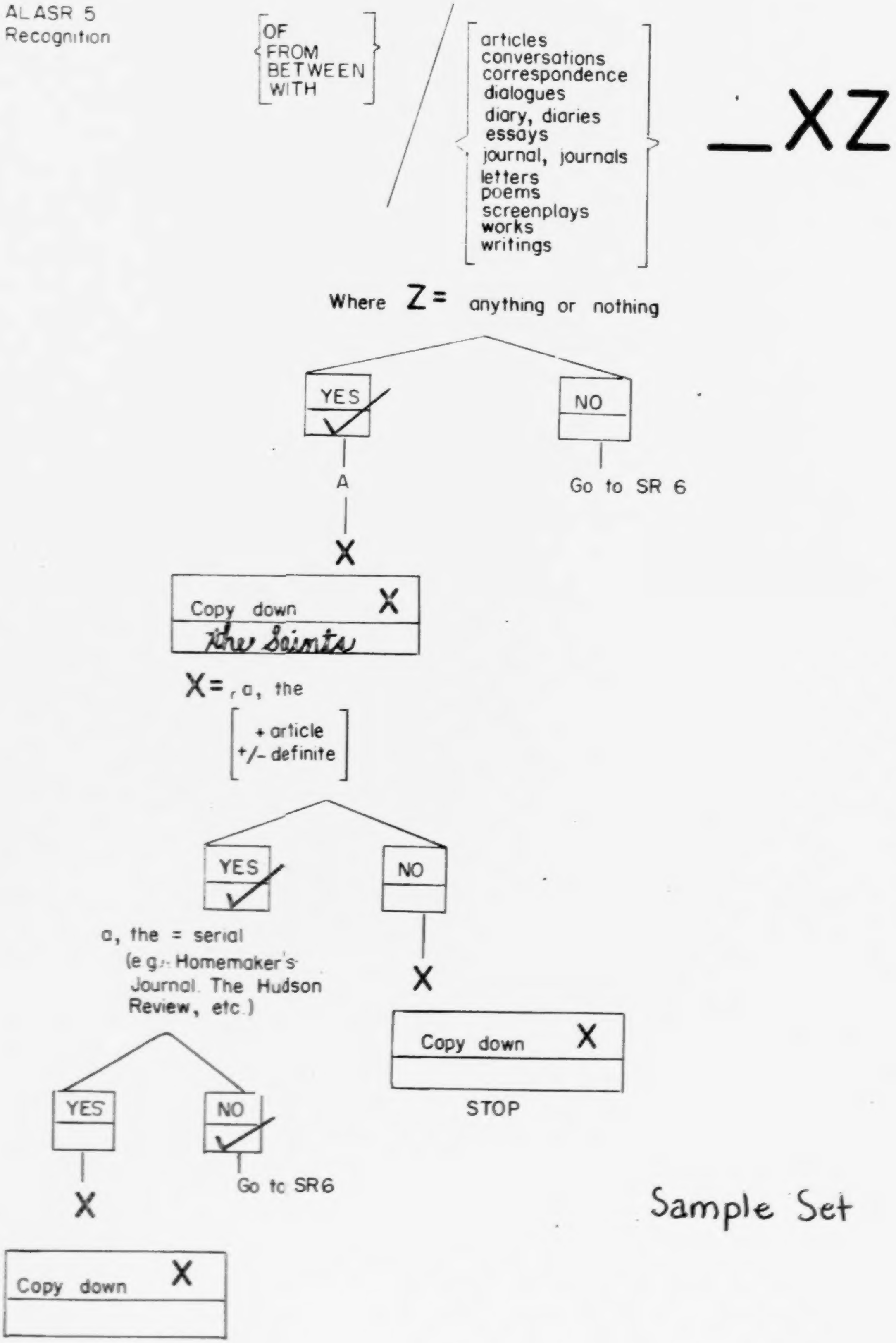
Where Z = anything or nothing



YES
NO
✓
A
Y
Go to SR 5
Copy down Y
Y= a, the
+article
+/-definite
YES
NO
a, the = serial
(e.g. A Homemaker's Journal, The Hudson Review, etc.)
Y
Copy down Y
STOP
YES
NO
Y
Go to SR 5
Copy down Y
STOP

Sample Set

ALASR 5
Recognition

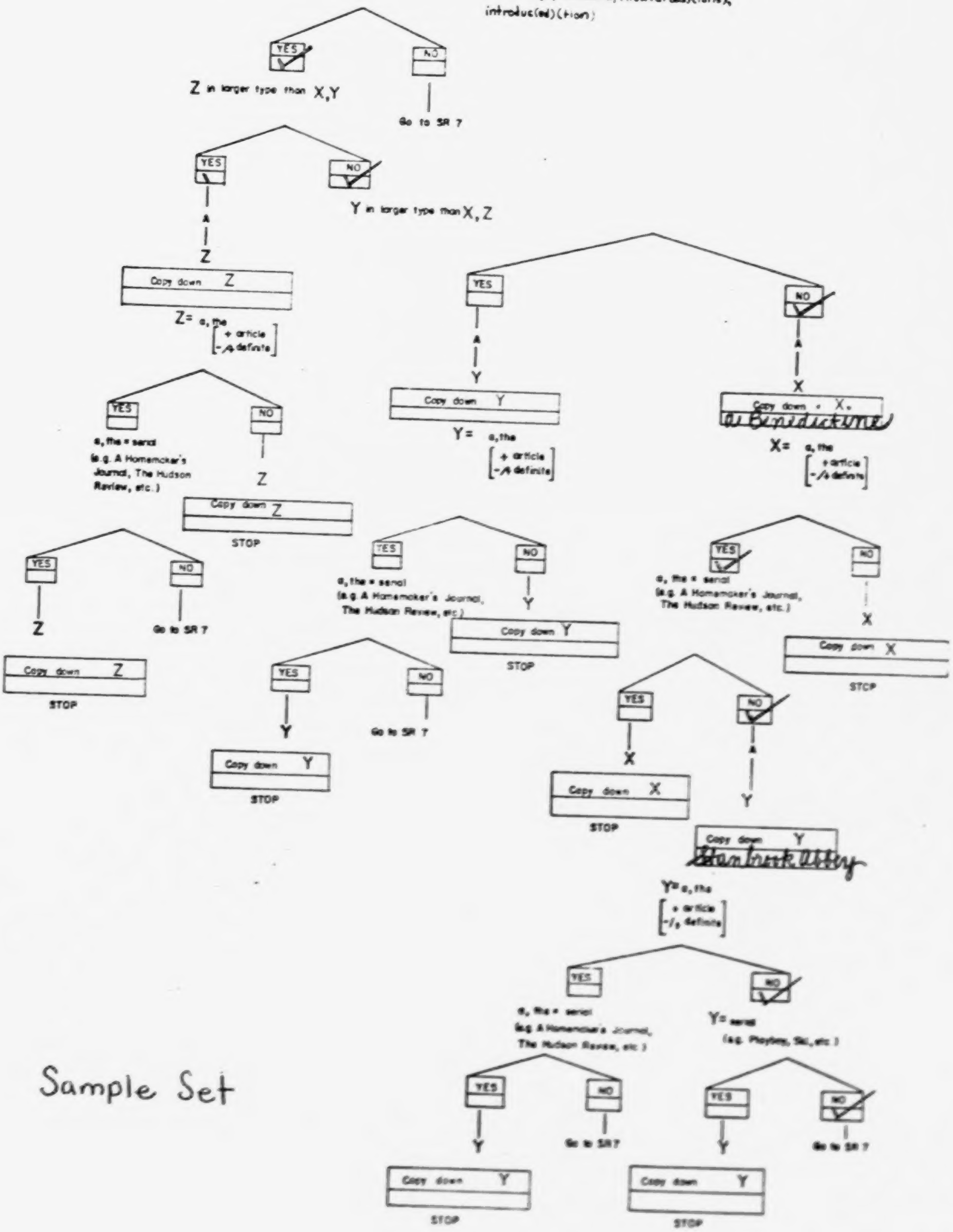

OF
FROM
BETWEEN
WITH
articles
conversations
correspondence
dialogues
diary, diaries
essays
journal, journals
letters
poems
screenplays
works
writings
_XZ
Where Z = anything or nothing
YES
✓
NO
A
Go to SR 6
X
Copy down X
the Saints
X = , a, the
+ article
+/- definite
YES
✓
NO
a, the = serial
(e.g., Homemaker's Journal. The Hudson Review, etc.)
X
Copy down X
STOP
YES
NO
✓
Go to SR6
X
Copy down X

Sample Set

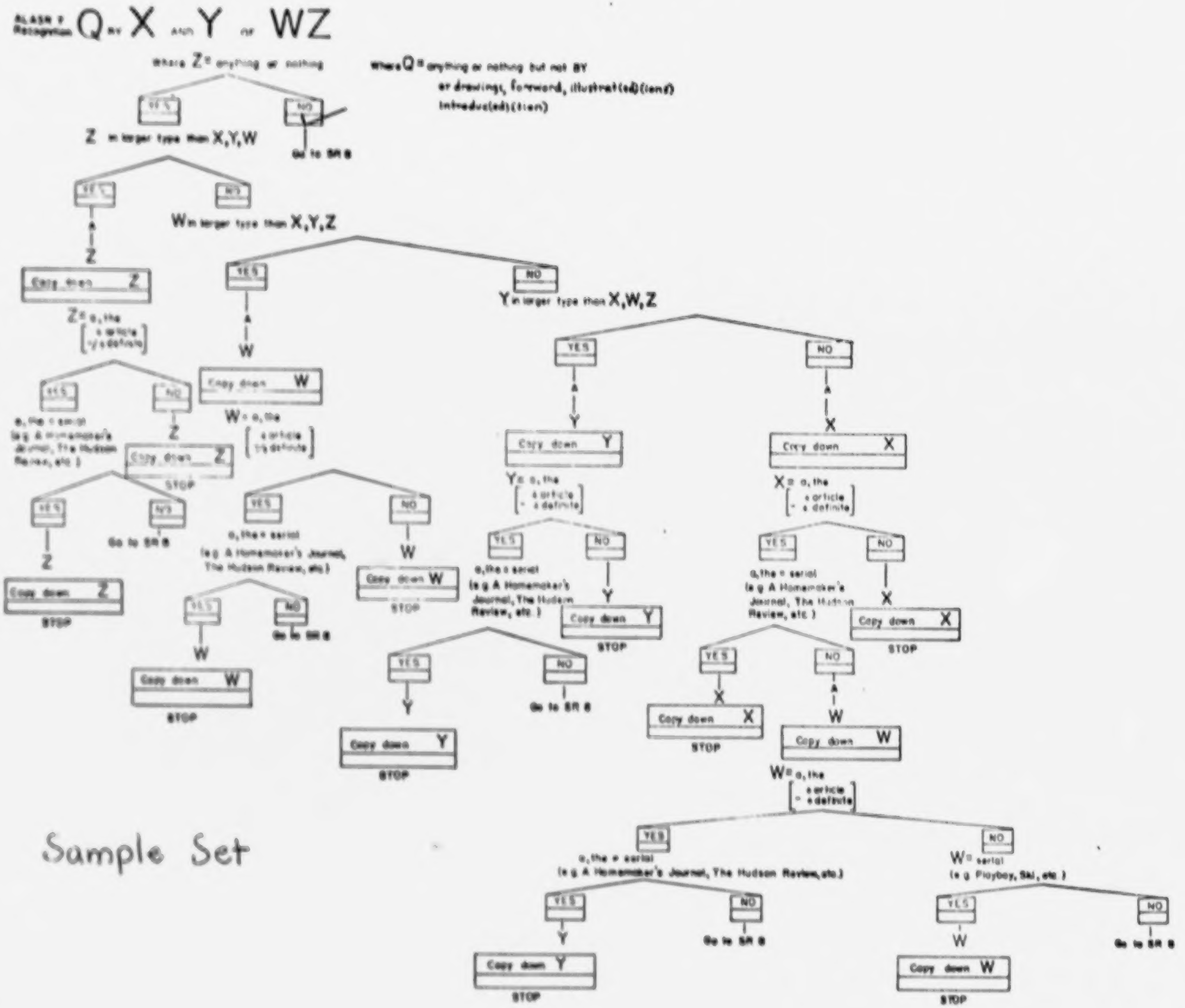
ALASR 6 Recognition Q BY X OF YZ

Where Z = anything or nothing

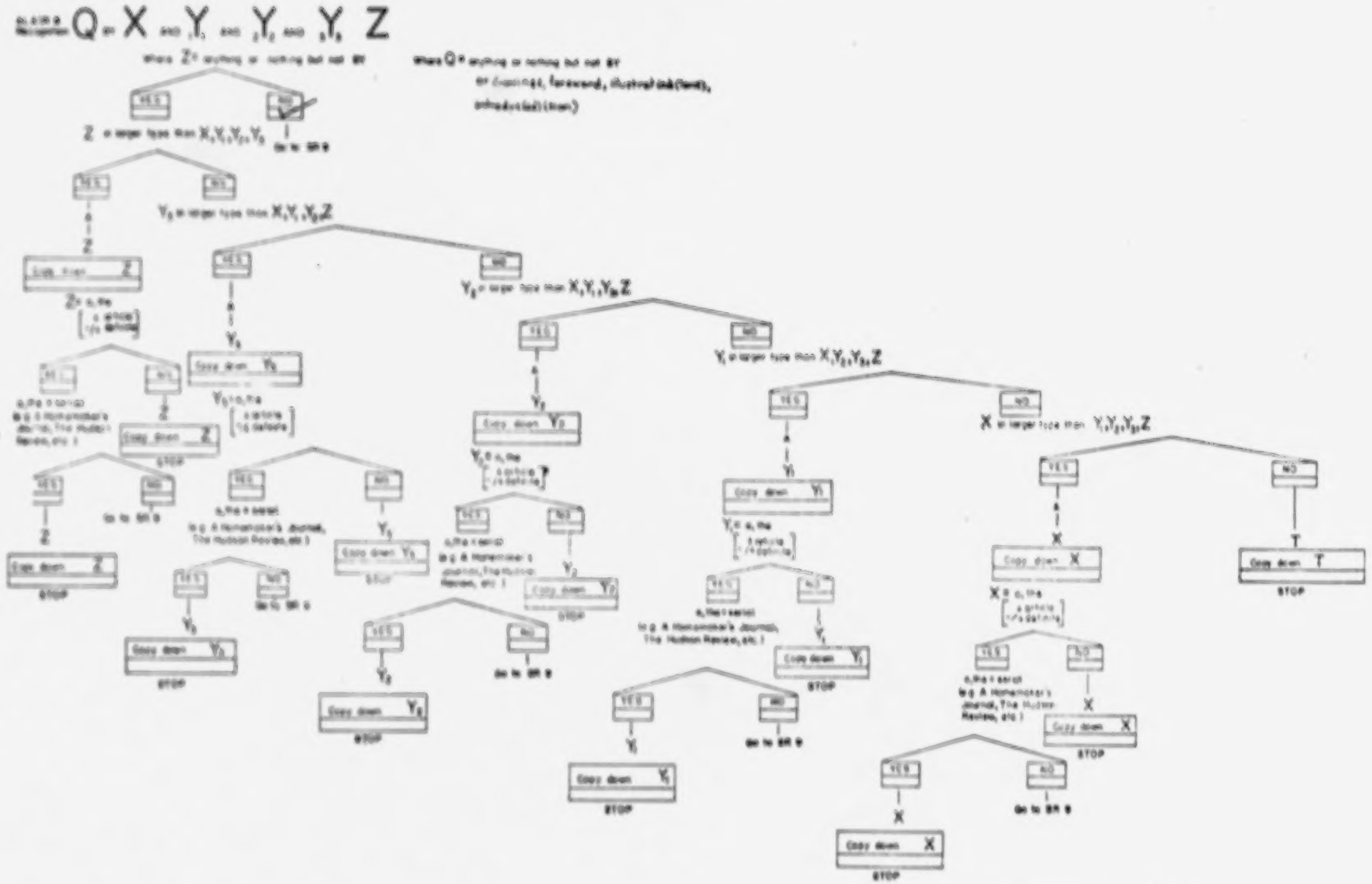
Where Q = anything or nothing but not BY or drawings, foreword, illustrat(ed)(ions), introduc(ed)(tion)


YES ✓
NO
Z in larger type than X,Y
Go to SR 7
YES
NO ✓
Y in larger type than X, Z
A
Z
Copy down Z
Z = a, the [+ article, -/+ definite]
YES
NO
a, the = serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)
Z
Copy down Z
STOP
YES
NO
Z
Copy down Z
STOP
Go to SR 7
YES
NO ✓
A
Y
Copy down Y
Y = a, the [+ article, -/+ definite]
A
X
Copy down X.
a Benedictine
X = a, the [+ article, -/+ definite]
YES
NO
a, the = serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)
Y
Copy down Y
STOP
YES
NO
Y
Copy down Y
STOP
Go to SR 7
YES ✓
NO
a, the = serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)
X
Copy down X
STOP
YES
NO ✓
X
Copy down X
STOP
A
Y
Copy down Y
Stanbrook Abbey
Y = a, the [+ article, -/+ definite]
YES
NO ✓
a, the = serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)
Y = serial (e.g. Playboy, Ski, etc.)
YES
NO
Y
Copy down Y
STOP
Go to SR 7
YES
NO ✓
Y
Copy down Y
STOP
Go to SR 7

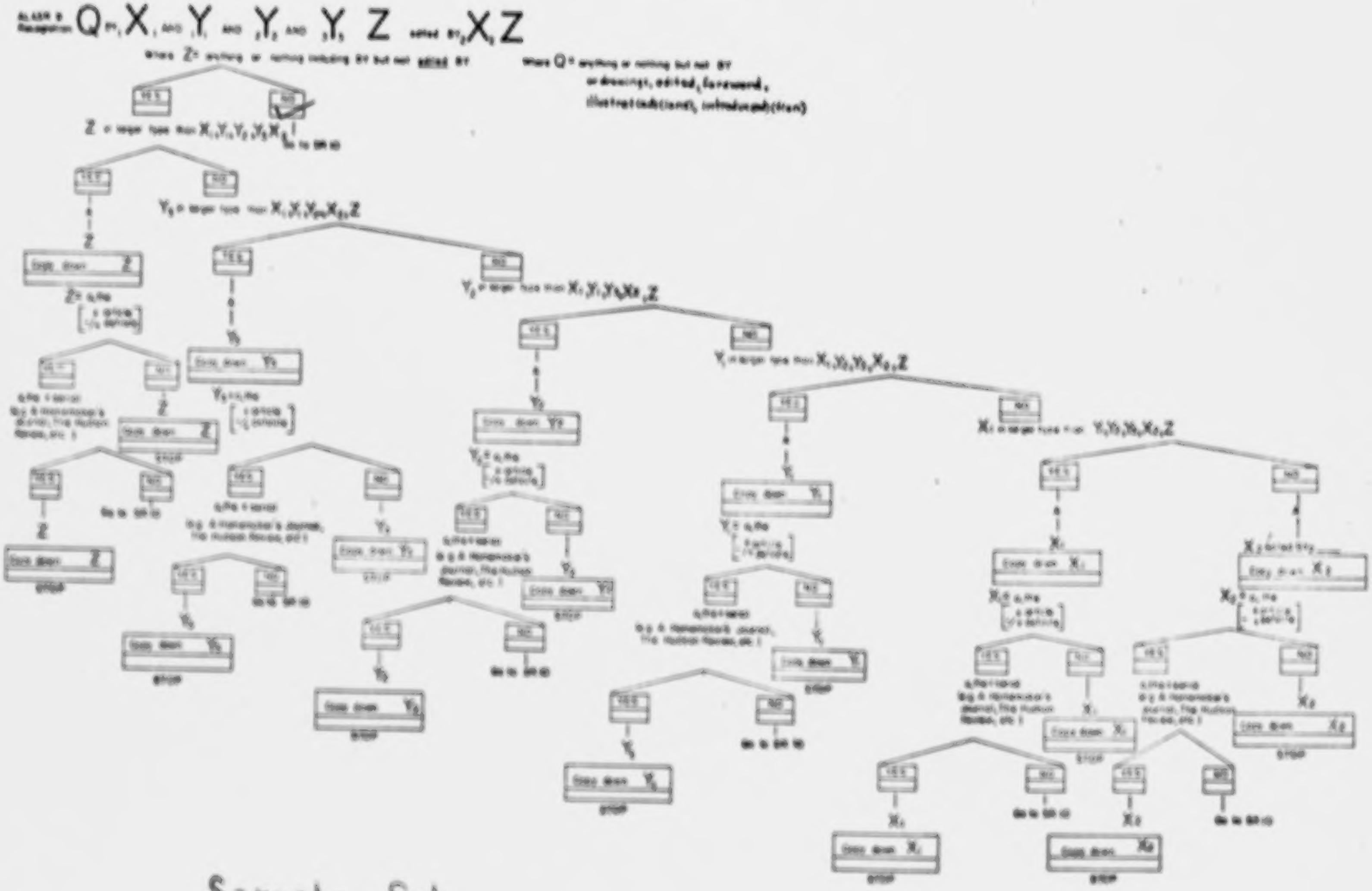
Sample Set


BLASR 7 Recognition Q BY X AND Y OF WZ
Where Z = anything or nothing
Where Q = anything or nothing but not BY
or drawings, foreword, illustrat(ed)(ions)
introduc(ed)(tion)
YES
NO
Go to SR 8
Z in larger type than X,Y,W
W in larger type than X,Y,Z
Y in larger type than X,W,Z
Copy down Z
Z = a, the [+ article +/- definite]
a, the + serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)
Copy down W
W = a, the [+ article +/- definite]
Copy down Y
Y = a, the [+ article + definite]
Copy down X
X = a, the [+ article + definite]
STOP
W = serial (e.g. Playboy, Ski, etc.)

Sample Set


Q X Y Y Y Z
YES
NO
Copy down Z
Copy down X
Copy down T
STOP

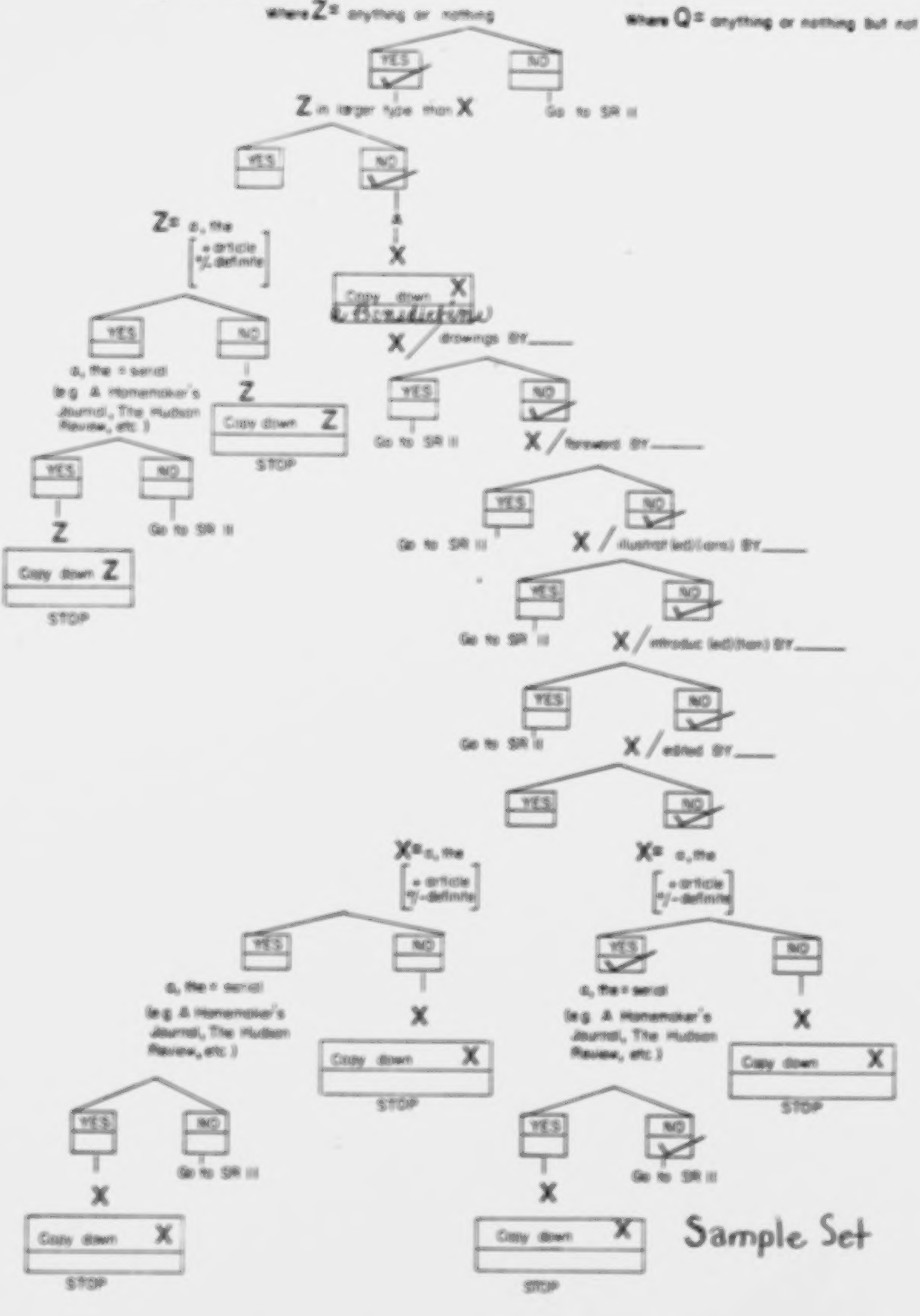
Sample Set


Q
X
Y
Y
Y
Z
X
Z
YES
NO
Copy down
STOP

Sample Set

Q BY XZ

Where Q = anything or nothing but not BY


Where Z = anything or nothing
YES ✓
NO
Z in larger type than X
Go to SR III
YES
NO ✓
a
X
Copy down X
a Benediction
X / drawings BY_____
YES
NO ✓
Go to SR III
X / foreword BY_____
YES
NO ✓
Go to SR III
X / illustrat(ed)(ions) BY_____
YES
NO ✓
Go to SR III
X / introduc(ed)(tion) BY_____
YES
NO ✓
Go to SR III
X / edited BY_____
YES
NO ✓
Z = a, the [+ article +/- definite]
YES
NO
a, the = serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)
Z
Copy down Z
STOP
YES
NO
Z
Copy down Z
STOP
Go to SR III
X = a, the [+ article +/- definite]
YES
NO
a, the = serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)
X
Copy down X
STOP
YES
NO
X
Copy down X
STOP
Go to SR III
X = a, the [+ article +/- definite]
YES ✓
NO
a, the = serial (e.g. A Homemaker's Journal, The Hudson Review, etc.)
X
Copy down X
STOP
YES
NO ✓
X
Copy down X
STOP
Go to SR III

Sample Set

ALASR II
Recognition

Everything else $\longrightarrow$ T

Where Everything else = recognitions to which previous rules did not apply

$\longrightarrow$ = goes to

T = Title

T

Copy down T

Letters from the Saints

STOP

Sample Set

APPENDIX E

TITLE PAGES USED IN THE EXPERIMENT

Start →

Ed McMahon's Barside Companion

Drawings by
Phil Interlandi ← Stop

The World Publishing Company
New York and Cleveland

417

AA b

Start → Left and Right with

LION AND RYAN

by Robert Littell

illustrated by Philip Wende ← Stop

COWLES BOOK COMPANY, INC.
NEW YORK

Start →

MALCOLM BOYD'S

Book of Days

← Stop

A FAWCETT CREST BOOK
Fawcett Publications, Inc., Greenwich, Conn.
Member of American Book Publishers Council, Inc.

Start → THE
SELECTED WORKS
OF
Cesare Pavese

TRANSLATED FROM THE ITALIAN
AND WITH AN INTRODUCTION BY
R. W. FLINT ← Stop

FARRAR, STRAUS AND GIROUX
NEW YORK

Start → SHEEP

IN RHODESIA

by a panel of experts and practical sheep farmers ← Stop

421
AA f

Start →

THE LETTERS OF
THOMAS WOLFE
TO HIS MOTHER

Newly edited from the original manuscripts

by C. HUGH HOLMAN
and SUE FIELDS ROSS

← Stop

Start →

THE OTHER SIDE

An Account of My Experiences with Psychic Phenomena

by

JAMES A. PIKE

with

DIANE KENNEDY ← Stop

Garden City, New York
DOUBLEDAY & COMPANY, INC.
1968

423
AA h

Start→

ANALYSIS AND SIMULATION OF MULTIPORT SYSTEMS

The Bond Graph Approach to Physical System Dynamics

by

Dean Karnopp and Ronald C. Rosenberg ←Stop

THE M.I.T. PRESS
Massachusetts Institute of Technology
Cambridge, Massachusetts, and London, England

424
AA 1

Start → De Staël-Du Pont Letters

Correspondence
of Madame de Staël and
Pierre Samuel du Pont de Nemours
And
of Other Members of the Necker
and du Pont Families

Edited and Translated by

JAMES F. MARSHALL ← Stop

The University of Wisconsin Press
Madison, Milwaukee, and London
1968

425
AA j

Start →

ROLES OF THE ATTORNEY GENERAL OF THE UNITED STATES

Luther A. Huston
Arthur Selwyn Miller
Samuel Krislov
Robert G. Dixon, Jr. ← Stop

July, 1968

PUBLISHED AND DISTRIBUTED BY THE
AMERICAN ENTERPRISE INSTITUTE
FOR PUBLIC POLICY RESEARCH
WASHINGTON, D.C. 20036

Start →

PASTORAL AND ROMANCE

MODERN ESSAYS IN CRITICISM

Edited by
Eleanor Terry Lincoln
Smith College ← Stop

PRENTICE-HALL, INC.
Englewood Cliffs, New Jersey

Start →

LISTEN TO A STORY

Kindergarten Stories and Verses
ILLUSTRATED BY CLAIRE SIMPSON

← Stop

ANGUS AND ROBERTSON

Start → SPOFFORD

BY *Herman Shumlin*

Adapted from the novel
Reuben, Reuben by Peter De Vries → Stop

RANDOM HOUSE · NEW YORK

429
AA n

Start →

MODERN EDUCATIONAL DANCE

BY
RUDOLF LABAN

SECOND EDITION
revised by
LISA ULLMANN
Director of the Laban Art of Movement Center ← Stop

FREDERICK A. PRAEGER, *Publishers*
New York · Washington

Start → BOTSFORD and ROBINSON'S

HELLENIC HISTORY

FIFTH EDITION

REVISED BY

DONALD KAGAN

PROFESSOR OF HISTORY, CORNELL UNIVERSITY ← Stop

THE MACMILLAN COMPANY
COLLIER-MACMILLAN LIMITED, LONDON

Start → better
scramble
than
lose

by FRAN TARKENTON
as told to JACK OLSEN ← Stop

FOUR WINDS PRESS/NEW YORK

Start → PLAYBOY'S
executive
reader

selected by the editors of Playboy ← Stop

PLAYBOY PRESS

Start → *The Life and Letters of*
VICE-ADMIRAL
LORD COLLINGWOOD

Oliver Warner ← Stop

LONDON
OXFORD UNIVERSITY PRESS
NEW YORK · TORONTO
1968

434

AA s

Start →

ASPECTS OF LANGUAGE AND LANGUAGE TEACHING

by W. A. BENNETT ← Stop

CAMBRIDGE
AT THE UNIVERSITY PRESS
1968

435
AA t

Start→LITTLE HENRY AND HIS BEARER

By
MRS. SHERWOOD

AND

THE LAST DAYS OF BOOSY

by
A. MacNEIL and THEOPHILUS SMITH ←Stop

HENRY E. WALTER LTD.
26 GRAFTON ROAD, WORTHING

Start → Reinterpretation in American Church History

BY R. PIERCE BEAVER, JERALD C. BRAUER
ROBERT T. HANDY, WINTHROP S. HUDSON
FREDERICK KIRSCHENMANN, MARTIN E. MARTY
SIDNEY E. MEAD, WILLIAM S. MORRIS

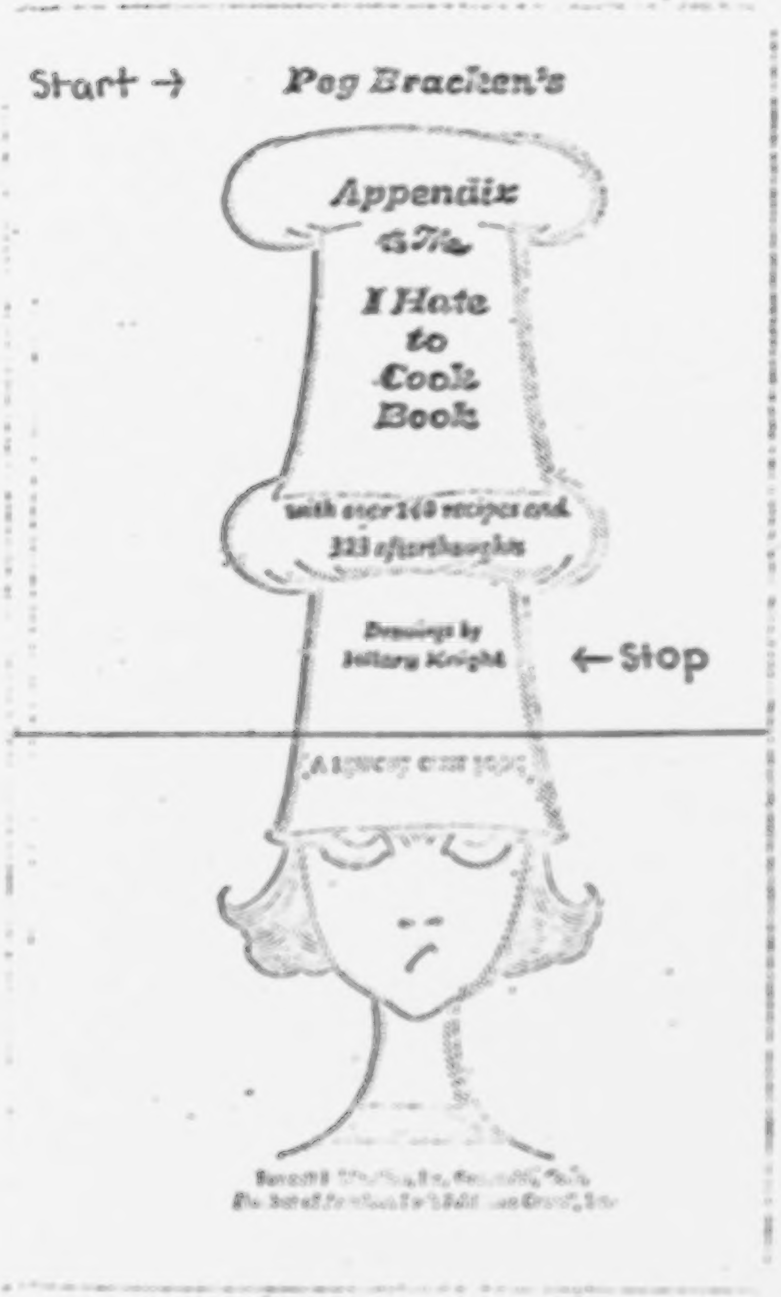
Edited by JERALD C. BRAUER ← Stop

THE UNIVERSITY OF CHICAGO PRESS
CHICAGO AND LONDON

437
AA v

Start → The
LINCOLN LIBRARY
of
SOCIAL STUDIES ← Stop

THE FRONTIER PRESS COMPANY
BUFFALO, NEW YORK 14203



Start →

Peg Bracken's

Appendix
to The
I Hate
to
Cook
Book

with over 140 recipes and
323 afterthoughts

Drawings by
Hilary Knight

← Stop

[illegible]

[illegible]
[illegible]

439

ALA B

Start →

NEIGHBORS IN SPACE

by

W. B. White

Illustrated by

Ruth C. Williams

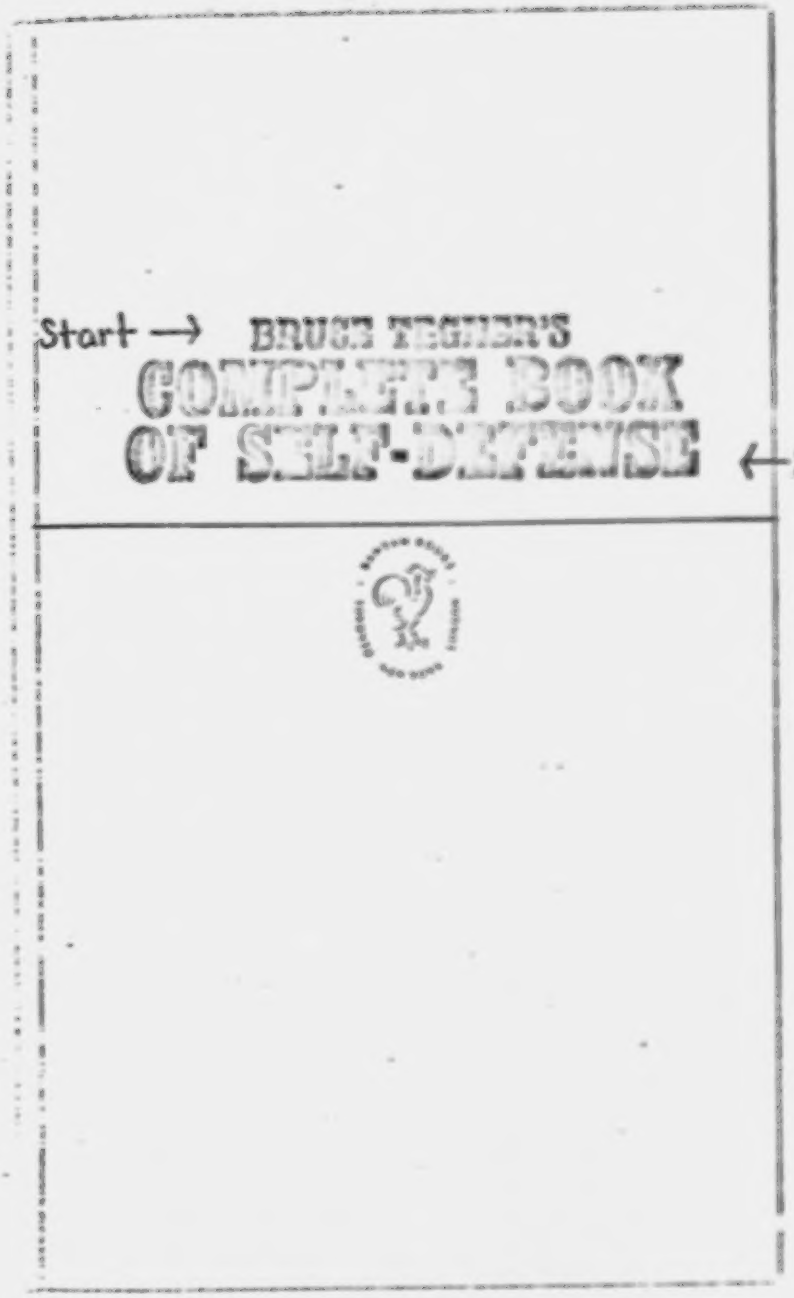
← Stop

Library of Congress Catalog Card Number: [illegible]

Copyright [illegible] as SEEING STARS by Rand McNally & Company
Copyright [illegible] as SEEING STARS
under International Copyright Union by Rand McNally & Company
All rights reserved

Printed in U.S.A. Edition of [illegible]

440
ALA C



Start → BRUCE TEGNER'S
COMPLETE BOOK
OF SELF-DEFENSE ← Stop

Start →

FOUR SCREENPLAYS
OF
INGMAR BERGMAN

Translated from the Swedish by LARS MALMSTROM and DAVID KUSHNER ← Stop

A CLARION BOOK
PUBLISHED BY SIMON AND SCHUSTER

441
ALA D

ALA E

Start →

HOW DID THE SATELLITES HAPPEN?

A Study of the Soviet Seizure
of Eastern Europe

by

A STUDENT OF AFFAIRS

with a Preface by

The Rt. Hon. HECTOR McNEIL, P.C., M.P.
(Sometime Minister of State for Foreign Affairs)

← Stop

LONDON
THE BATCHWORTH PRESS

443

ALA F

Start → *The Letters of*

JOHN RUSKIN

to Lord and Lady Mount-Temple

Edited and with an Introduction by

John Lewis Bradley ← Stop

Ohio
State
University
Press

ALA G

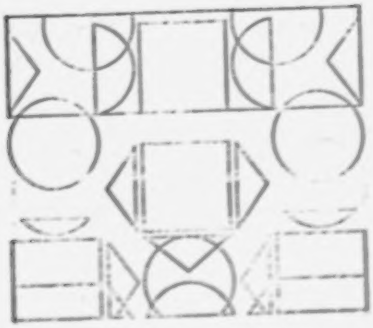
Start → SOMETIMES I WONDER

The Story of Hoagy Carmichael

by Hoagy Carmichael
with
Stephen Longstreet ← Stop

FARRAR, STRAUS AND GIROUX
NEW YORK

ALA H



Start→ THE MATHEMATICAL THEORY OF COMMUNICATION

by Claude E. Shannon and Warren Weaver ←Stop

THE UNIVERSITY OF ILLINOIS PRESS · URBANA · CHICAGO · LONDON

Start→ The Dialogues of Archibald MacLeish and Mark Van Doren

EDITED BY

WARREN V. BUSH ←Stop

NEW YORK ◊ E. P. DUTTON & CO., INC. ◊ 1964

Start →

CURRENT TRENDS IN THE DESCRIPTION AND ANALYSIS OF BEHAVIOR

ROBERT GLASER
JOSEPH ZUBIN
DONALD B. LINDSLEY
VINCENT NOWLIS
CHARLES N. COFER
HAROLD GUETZKOW
JOHN B. CARROLL
ROY M. HAMLIN
THOMAS M. FRENCH ← Stop

1958
UNIVERSITY OF PITTSBURGH PRESS

448
ALA K

Start → *Essays on the*

Eighteenth-Century Novel

Edited by

ROBERT DONALD SPECTOR ← Stop

INDIANA UNIVERSITY PRESS
Bloomington and London

Start →

A TREASURE CHEST OF POETRY

⁂

ILLUSTRATED
BY
LASZLO
MATULAY

← Stop

PARENTS' MAGAZINE PRESS
A Division of Parents' Magazine Enterprises, Inc., New York

Start → *The Wall*

A DRAMA IN THREE ACTS

By Millard Lampell

Based on the novel by John Hersey ← Stop

SAMUEL FRENCH, INC.
25 West 45th Street New York, N. Y., 10036
7623 Sunset Boulevard Hollywood, Calif., 90046
LONDON *TORONTO*

Start → *The Standard*

BARTENDER'S GUIDE

BY

PATRICK GAVIN DUFFY

•

Revised and enlarged by

JAMES A. BEARD ← Stop

PERMABOOKS • NEW YORK

452
ALA 0

Start→TOPLEY AND WILSON'S PRINCIPLES OF BACTERIOLOGY AND IMMUNITY

FOURTH EDITION

By

G. S. WILSON, M.D., F.R.C.P., D.P.H.

DIRECTOR OF THE PUBLIC HEALTH LABORATORY SERVICE,
ENGLAND AND WALES

AND

A. A. MILES, C.B.E., M.A., M.D., F.R.C.P.

PROFESSOR OF EXPERIMENTAL PATHOLOGY, UNIVERSITY OF LONDON,
AND DIRECTOR OF THE LISTER INSTITUTE OF PREVENTIVE MEDICINE, LONDON ←Stop

---

IN TWO VOLUMES
VOLUME I

LONDON
EDWARD ARNOLD (PUBLISHERS) LTD.

Start → DUDE WRANGLER
:
HUNTER
:
LINE RIDER

by

FLOYD C. BARD
as told to
AGNES WRIGHT SPRING ← Stop

Sage Books
Denver

Start→ SKI MAGAZINE'S

SKI POINTERS
by the Experts

BY THE EDITORS OF *SKI MAGAZINE*

Illustrated ←Stop

Award Books • New York City

456

ALA R

Start → THE INEVITABLE GUEST

Life and Letters of Jemima Darby

by

JOHN ARDIS CAWTHON ← Stop

The Naylor Company
Book Publishers of the Southwest
San Antonio, Texas

Start → THE STRUCTURE OF SCIENTIFIC REVOLUTIONS

By THOMAS S. KUHN ← Stop

THE UNIVERSITY OF CHICAGO PRESS
CHICAGO & LONDON

457
ALA 7

HARVARD STUDIES AND NOTES
IN PHILOLOGY AND LITERATURE

VOLUME XII

Start→THE DON CARLOS THEME

By FREDERICK W. C. LIEDER

THE RUSSIAN PRIMARY CHRONICLE

By SAMUEL H. CROSS ←Stop

CAMBRIDGE
HARVARD UNIVERSITY PRESS
1930

458
ALA 0

Start→ Actor and Architect

by

TYRONE GUTHRIE
RICHARD SOUTHERN
SEAN KENNY
CHRISTOPHER STEVENS
HUGH HUNT
JOHN ENGLISH

edited by

STEPHEN JOSEPH ←Stop

UNIVERSITY OF TORONTO PRESS

459

ALA V

Start THE AMERICAN PEOPLES ENCYCLOPEDIA

A comprehensive reference work Stop

GROLIER INCORPORATED • NEW YORK

APPENDIX F

QUESTIONNAIRE TO PARTICIPANTS AND CORRESPONDENCE

POST-TEST QUESTIONNAIRE

1. Sets of title pages were ordered alphabetically. Did you

   1 Apply title pages to the rules in their alphabetical order

   2 Employ random selection (selecting those that seemed easier to do first, etc.)

2. A set of recognition rules (RR) and selection rules (SR) were applied to each title page. Would you say

   1 The RR were easier to apply than the SR

   2 The SR were easier to apply than the RR

   3 There was no difference in the degree of ease in applying either set of rules

   4 I do no remember

   Comments, if any, on the response you circled ______________

   ____

   ____

3. Two sets of recognition rules (RR) and selection rules (SR) were used in the experiment. Half of the title pages required use of a set identified as AA RR and SR while the other half required use of a set identified as ALA RR and SR. From what you recall, would you say

   1 The AA RR and SR were easier to apply than the ALA RR and SR

   2 The ALA RR and SR were easier to apply than the AA RR and SR

   3 There was no difference in the degree of ease in applying either set of rules

   4 I do not remember

   Comments, if any, on the response you circled ______________

   ____

   ____

4 Taking the rules as a whole, would you say that after their application to the first few title pages, they were

1 Extremely mechanical

2 Semi-mechanical

3 Mechanical to a large degree

4 Mechanical to a small degree

5 Not mechanical at all

Comments, if any, on the response you circled ______________

____

____

THE LIBRARY OF CONGRESS

Washington, D. C. 20540

Processing Department
Office Of The Director

October 8, 1969

Dear Mrs. Fox:

We were interested to learn of your doctoral study involving a comparison of form and choice of entry under the ALA and AA cataloging rules. We regret, however, that the Library does not keep statistics that shed light on the frequency of use of particular rules nor can we readily indicate the frequency of particular types of entry.

It may be possible to provide a partial answer to the second question by analyzing the growing file of MARC records but there are two reasons why such evidence would have to be treated with great circumspection. First, a MARC record does not indicate which set of rules was used to establish the headings it contains. Second, the present file contains only records for English-language monographs and thus is not a valid representation of the totality of LC cataloging output. A study of the frequency of data elements in the MARC file is now under way and it is hoped that the results can be published in the next several months. Despite the limitations of the data for your purposes, you may find them of some use when they are available.

Sincerely yours,

John C. Rather
Specialist in Technical
Processes Research

VITA

Ann Sandberg Fox was born in Providence, Rhode Island. She attended elementary and secondary schools in Norwich and Baltic, Connecticut. She received her AB degree in English literature from the College of Our Lady of the Elms in 1951. She received her master's degree in library science from Catholic University of America, in 1955. In 1958 she attended Trinity College, Dublin, Ireland, where she was enrolled in the Department of Celtic Studies. She was elected to Delta Epsilon Sigma, National Catholic Honor Society, and Beta Phi Mu, International Library Science Honor Society, and was the recipient of numerous academic awards and fellowships. While pursuing her doctoral degree, she was awarded an HEA fellowship, and was selected as a University of Illinois fellow. Her doctoral dissertation was supported by a U.S.O.E. Small Project Research grant.

Mrs. Fox has wide professional experience. She held the position of supervisory librarian in Germany with the U.S. Special Services Section from 1955 to 1957. She was an English and Celtic language cataloger in the Descriptive Cataloging Division of the Library of Congress in 1958, and was appointed head of the Cataloging Unit of the Preliminary Cataloging Section in that Division in 1959. Before entering the doctoral program, she was a visiting lecturer in the Graduate School of Library Science at the University of Illinois in the summer of 1964. While enrolled in the doctoral program, she was appointed an instructor in the program, teaching a cataloging course on campus as well as on extension.

Mrs. Fox has had poems published in two anthologies, and an extensive set of visuals that she prepared in teaching are presently being edited for publication.